# Ralph Chill, Eva Fašangová 

## Gradient Systems

- 13th International Internet Seminar -

February 10, 2010
(c) by R. Chill, E. Fašangová

## Contents

0 Introduction ..... 1
1 Gradient systems in euclidean space ..... 5
1.1 The space $\mathbb{R}^{d}$ ..... 5
1.2 Fréchet derivative ..... 7
1.3 Euclidean gradient. ..... 8
1.4 Euclidean gradient systems ..... 9
1.5 Algebraic equations and steepest descent ..... 12
1.6 Exercises ..... 13
2 Gradient systems in finite dimensional space ..... 15
2.1 Definition of gradient ..... 17
2.2 Definition of gradient system ..... 18
2.3 Local existence for ordinary differential equations ..... 18
2.4 Local and global existence for gradient systems ..... 21
2.5 Newton's method ..... 23
2.6 Exercises ..... 25
3 Gradients in infinite dimensional space:
the Laplace operator on a bounded interval ..... 27
3.1 Definition of gradient ..... 28
3.2 Gradients of quadratic forms ..... 30
3.3 Sobolev spaces in one space dimension ..... 31
3.4 The Dirichlet-Laplace operator on a bounded interval ..... 33
3.5 The Dirichlet-Laplace operator with multiplicative coefficient ..... 35
3.6 Exercises ..... 37
4 Gradients in infinite dimensional space: the Laplace operator and the $p$-Laplace operator ..... 39
4.1 Sobolev spaces in higher dimensions ..... 39
4.2 The Dirichlet-Laplace operator ..... 42
4.3 The Dirichlet- $p$-Laplace operator ..... 44
4.4 An auxiliary result ..... 46
4.5 Origins of the Laplace operator ..... 49
4.6 Exercises ..... 52
5 Bochner-Lebesgue and Bochner-Sobolev spaces ..... 55
5.1 The Bochner integral ..... 56
5.2 Bochner-Lebesgue spaces ..... 58
5.3 Bochner-Sobolev spaces in one space dimension ..... 59
5.4 Exercises ..... 63
6 Gradient systems in infinite dimensional spaces: existence and uniqueness of solutions ..... 65
6.1 Gradient systems in infinite dimensional spaces ..... 65
6.2 Global existence and uniqueness of solutions for gradient systems with convex energy ..... 66
6.3 Exercises ..... 75
7 Diffusion equations ..... 77
7.1 Heat conduction ..... 77
7.2 The reaction-diffusion equation ..... 80
7.3 The Perona-Malik model in image processing ..... 81
7.4 Global existence and uniqueness of solutions ..... 82
7.5 Exercises ..... 84
8 Gradient systems in infinite dimensional spaces: existence and uniqueness of solutions II ..... 87
8.1 Global existence and uniqueness of solutions for gradient systems with elliptic energy ..... 88
8.2 Exercises ..... 97
9 Regularity of solutions ..... 99
9.1 Interpolation ..... 99
9.2 Time regularity - Angenent's trick ..... 101
9.3 Exercises ..... 104
10 Existence of local solutions ..... 107
10.1 The local inverse theorem and existence of local solutions of abstract differential equations ..... 108
10.2 Gradients of quadratic forms and identification of the trace space. ..... 109
10.3 A semilinear diffusion equation ..... 112
10.4 Exercises ..... 117
11 Asymptotic behaviour of solutions of gradient systems ..... 119
11.1 The $\omega$-limit set of a continuous function on $\mathbb{R}_{+}$ ..... 120
11.2 A stabilisation result for global solutions of gradient systems ..... 121
11.3 Nonstabilisation results for global solutions of gradient systems ..... 123
11.4 Exercises ..... 126
12 Asymptotic behaviour of solutions of gradient systems II ..... 129
12.1 The Łojasiewicz-Simon inequality and stabilisation of global solutions of gradient systems ..... 130
12.2 The Łojasiewicz-Simon inequality in Hilbert spaces ..... 133
12.3 Exercises ..... 138
13 The universe, soap bubbles and curve shortening ..... 141
References ..... 151
A Primer on topology ..... 1001
A. 1 Metric spaces ..... 1001
A. 2 Sequences, convergence ..... 1004
A. 3 Compactness ..... 1006
A. 4 Continuity ..... 1007
A. 5 Completion of a metric space ..... 1008
B Banach spaces and bounded linear operators ..... 1011
B. 1 Normed spaces ..... 1011
B. 2 Product spaces and quotient spaces ..... 1017
B. 3 Bounded linear operators ..... 1020
B. 4 The Arzela-Ascoli theorem ..... 1025
C Calculus on Banach spaces ..... 1029
C. 1 Differentiable functions between Banach spaces ..... 1029
C. 2 Local inverse theorem and implicit function theorem ..... 1030
C. 3 * Newton's method ..... 1034
D Hilbert spaces ..... 1035
D. 1 Inner product spaces ..... 1035
D. 2 Orthogonal decomposition ..... 1039
D. 3 * Fourier series ..... 1044
D. 4 Linear functionals on Hilbert spaces ..... 1048
D. 5 Weak convergence in Hilbert spaces ..... 1049
E Dual spaces and weak convergence ..... 1053
E. 1 The theorem of Hahn-Banach ..... 1053
E. 2 Weak* convergence and the theorem of Banach-Alaoglu ..... 1059
E. 3 Weak convergence and reflexivity ..... 1060
E. 4 * Minimization of convex functionals ..... 1065
E. 5 * The von Neumann minimax theorem ..... 1068
F The divergence theorem (By Wolfgang Arendt) ..... 1071
F. 1 Open sets of class $C^{1}$ ..... 1071
F. 2 The divergence theorem ..... 1074
F. 3 Proof of the divergence theorem ..... 1077
G Sobolev spaces ..... 1083
G. 1 Test functions, convolution and regularization ..... 1083
G. 2 Sobolev spaces in one dimension ..... 1086
G. 3 Sobolev spaces in several dimensions ..... 1092
G. 4 * Elliptic partial differential equations ..... 1094
H Bochner-Lebesgue and Bochner-Sobolev spaces ..... 1097
H. 1 The Bochner integral ..... 1097
H. 2 Bochner-Lebesgue spaces ..... 1100
H. 3 Bochner-Sobolev spaces in one dimension ..... 1102
Index ..... 1105

## Lecture 0 <br> Introduction

Philosophy is written in that gigantic book which is perpetually open in front of our eyes (I allude to the universe), but no one can understand it who does not strive beforehand to learn the language and recognize the letters in which it is written. It is written in mathematical language, and its letters are triangles, circles and other geometrical figures, and without these means it is humanly impossible to understand any of it; without them, all we can do is to wander aimlessly in an obscure labyrinth.

## Galileo Galilei

The evolution of many systems from physics, biology, chemistry, engineering, economics, social sciences can be rewritten in the mathematical language which uses ordinary or partial differential equations. Due to the analysis of the differential equations we have the hope that we need not wander aimlessly in an obscure labyrinth but that we understand the corresponding evolution problems.

In order to illustrate this statement, think of the motion of a pendulum, for example the oscillations of the great lantern in the nave of Pisa cathedral. At the end of the 16th century, after mere observation of this lantern and probably other pendula, Galilei conjectured that on the one hand the period of one swing is proportional to the length of the pendulum and on the other hand this period is independent of the amplitude of the swing. The conjecture about the precise period of one swing would have many consequences if it was true, especially since it is connected to the problem of measuring time. But how could Galilei confirm his conjecture? Could or can it be confirmed mathematically? Discovered some 30 years after Galilei by Newton and Leibniz, the differential calculus (the calculus of fluxions, as it was called by Newton) turned out to be an appropriate mathematical language into which the oscillations of the great lantern in the nave of Pisa cathedral could be translated. At least then, by the mathematical analysis of the corresponding differential equation, there was an explanation why Galilei's conjecture was very close to the truth but not quite the truth; in particular, there was an explanation why the period of one swing does depend on the amplitude. Moreover, the differential calculus could be used to construct noncircular pendula for which the period does not depend on the amplitude, that is, to solve the so-called
tautochrone problem ${ }^{1}$. But the success of the new language of differential calculus may be illustrated by a second classical example: the motion of planets and stars. The problem to describe and predict their motions fascinated astronomers since the ancient times, but there was, like in the case of the pendulum, a language missing to explain all the observations of the great astronomers. Starting from his law of gravitation and his second law of equilibrium of forces, and by translating these laws into the mathematical language of differential equations, Newton was able to confirm Kepler's laws about the elliptic orbits and the regular motions of the planets in our solar system.

Today, more than three centuries after Newton wrote down his treatise on the theory of fluxions, and after the rise of the calculus of variations, it has turned out that a large number of evolution models may be expressed in the mathematical language of ordinary and partial differential equations. Besides the evolution models from classical mechanics, the mechanics of points and solids, we want to mention other phenomena such as the evolution of heat or diffusing particles, the evolution of waves (water waves, electromagnetic waves, the oscillations of elastic materials), the evolution of cells or other populations, the evolution of soap bubbles or other surfaces. These are only a few, however important examples besides many others. Moreover, it is natural to expect that the list of all phenomena which may be expressed in the mathematical language using ordinary and partial differential equations is still open. For all the phenomena, one aspect, which is mentioned in Galilei's citation, is important for our work: the mathematical analysis of the various differential equations and the interpretations of the qualitative properties which we deduce from this analysis allow us to understand the various evolution phenomena in a deeper way. One aim of this course is to make a small step in this process of understanding, as far as gradient systems are concerned. We believe that the class of gradient systems forms a fundamental class within the differential equations.

The translation of evolution models into the mathematical language using ordinary or partial differential equations very often starts from first principles like the equilibrium of forces, the conservation of total energy, the conservation of total mass or the conservation of total number of individuals. One may therefore be surprised when discovering that the resulting differential equations are dissipative in the sense that a characteristic quantity - like for example again some form of energy - is strictly decreasing in time, unless the system is at rest. Although in the theory of dynamical systems the word dissipative is used in a wider sense, we think it reflects well the existence of a strictly decreasing quantity: the Concise Oxford Dictionary translates the verb dissipate by disperse, dispel, (cause to) disappear, (cloud, vapour, care, fear, darkness); break up entirely, bring or come

[^0]to nothing; squander (money); fritter away (energy, attention). Galilei's lantern in Pisa, like any other mechanical pendulum, stabilizes and finally comes to rest because of inner friction forces and because of the friction between the lantern and the surrounding air. The friction forces, however small they are compared to the leading gravitational force, can not be completely neglected when translating Newton's second law into a second order differential equation. The analysis of this differential equation shows that the friction is the reason for the dissipation of the total energy of the pendulum (kinetic energy plus potential energy). For a similar reason, a vibrating string on a guitar or a violon eventually ceases to vibrate if let alone, and the same is true for a drum or any other vibrating plate.

The simplest partial differential equation arising from the problem of heat conduction is the linear heat equation

$$
u_{t}-\Delta u=0
$$

It is a result of the principle of conservation of total heat and of Fourier's law, as we will see in this course. At the same time we shall see that the linear heat equation is a dissipative evolution equation. Both properties, conservation (of heat) and dissipation (of energy), are not necessarily contradictory: experiments confirm that due to heat conduction in a physical material the heat distribution in an isolated volume eventually tends to a uniform, constant distribution. In other words, there is an effect of stabilization of the heat distribution, while the total amount of heat remains conserved. The effect of stabilization can be explained by the existence of a characteristic quantity which is dissipated by the heat equation, and it is only a side-order remark that in the case of the linear heat equation there are actually infinitely many such quantities. But even more is true, as we shall see: the linear heat equation, as well as the more general diffusion equation

$$
u_{t}-\operatorname{div} c(x,|\nabla u|) \nabla u=0,
$$

the Cahn-Hilliard equation from phase separation models, the mean curvature equation, and many other dissipative parabolic evolution equations can be written as abstract gradient systems.

The abstract gradient system

$$
\dot{u}+\nabla \mathscr{E}(u)=0
$$

is the prototype example of a dissipative evolution equation. In this abstract differential equation, $\mathscr{E}$ is a real-valued function defined on an open subset of $\mathbb{R}^{d}$, of a Banach space, or of a finite or infinite dimensional manifold. The function $\mathscr{E}$ plays the role of the characteristic quantity mentioned above. Its gradient $\nabla \mathscr{E}$ depends on the ambient metric and will be defined later; for the moment being it suffices to think of the case that $\mathscr{E}$ is defined on $\mathbb{R}^{d}$ and $\nabla \mathscr{E}$ is the usual euclidean
gradient: $\nabla \mathscr{E}(u)=\left(\frac{\partial \mathscr{E}}{\partial u_{1}}(u), \ldots, \frac{\partial \mathscr{E}}{\partial u_{d}}(u)\right)$.
Let $u$ be a solution of the above gradient system, that is, $u$ is a differentiable $\mathbb{R}^{d}$-valued function defined on an interval $I \subseteq \mathbb{R}$ which satisfies the equation $\dot{u}+$ $\nabla \mathscr{E}(u)=0$ everywhere on $I$. If $\mathscr{E}(u)$ denotes the composition of the functions $\mathscr{E}$ and $u$, then one calculates

$$
\begin{aligned}
\frac{d}{d t} \mathscr{E}(u(t)) & =\langle\nabla \mathscr{E}(u(t)), \dot{u}(t)\rangle \\
& =-\langle\dot{u}(t), \dot{u}(t)\rangle \leq 0
\end{aligned}
$$

here we denoted by $\langle\cdot, \cdot\rangle$ the euclidean inner product, we used the chain rule and the assumption that $u$ is a solution. This simple but fundamental relation shows that the function $\mathscr{E}(u)$ is strictly decreasing in time unless the solution $u$ is constant, which would mean that the underlying evolution is at rest. In other words, the quantity $\mathscr{E}$, which may be some form of energy, is dissipated by the gradient system. This fundamental property of a gradient system, namely dissipativity, but also the gradient structure itself may be used to prove existence and uniqueness of solutions, and to obtain qualitative results about their regularity and their long-time behaviour, like the stabilization or non-stabilization of solutions.

Let us end this Introduction by a philosophical question. The dissipation of some form of energy is a characteristic property of many evolution phenomena, and many results which will be proved in this course remain true in a more general context of dissipative evolution equations; we see this as a motivation for our approach, too. But at the same time some questions arise: how much larger is the class of dissipative evolution equations compared to the class of gradient systems? How to show that some dissipative ordinary or partial differential equation is a gradient system? And how to show that it is not a gradient system? Very probably we will not be able to answer these farreaching questions which are already interesting for the Navier-Stokes equations. Nevertheless, by modifying a well known citation from Einar Hille we would like to say: we hail a gradient system whenever we see one and we seem to see them everywhere ${ }^{2}$... We hope that the reader of this course will share our impression.

[^1]
## Lecture 1

## Gradient systems in euclidean space

We start by recalling the definition of the derivative of a real-valued function on $\mathbb{R}^{d}$, we define the euclidean gradient of such a function, we define what we mean by a euclidean gradient system, and we present an example of such a gradient system.

We believe that the definition of derivative and euclidean gradient should be well known for the reader, and we shall in the sequel use several results from differential calculus, like for example the chain rule, the product rule, and also the definition of the derivative of a function between finite (or infinite) dimensional spaces without proving or even stating them (some results, however, can be found in the Appendix). By recalling the definition of derivative and gradient, despite of our believe that they are well known, we only want to make sure from the very beginning that they are different objects. Let us say it in this form: a function comes always along with its derivative, while this is not immediately the case for a gradient which depends in addition on the geometry of the ambient space on which the function is defined.

First properties of the gradient system

$$
\dot{u}+\nabla_{e u c} \mathscr{E}(u)=0
$$

are studied. We introduce the concept of energy for a general ordinary differential equation, that is, a quantity which is decreasing along solutions of gradient systems. The phenomenon of a decreasing quantity is very frequent in many examples of ordinary differential equations which the nature provides to us, and it is this phenomenon and its consequences which are discussed in several places of this course.

### 1.1 The space $\mathbb{R}^{d}$

We fix some notation in $\mathbb{R}^{d}$. Elements of $\mathbb{R}^{d}$ are denoted in one of the following ways:

$$
u \text { or }\left(u_{i}\right) \text { or }\left(u_{i}\right)_{1 \leq i \leq d} \text { or }\left(u_{1}, \ldots, u_{d}\right) \text { or }\left(\begin{array}{c}
u_{1} \\
\vdots \\
u_{d}
\end{array}\right)
$$

in particular, we do not distinguish between column vectors and row vectors.
The euclidean inner product on $\mathbb{R}^{d}$ is given by

$$
\langle u, v\rangle=\langle u, v\rangle_{e u c}=\sum_{i=1}^{d} u_{i} v_{i} \quad \text { for every } u=\left(u_{i}\right), v=\left(v_{i}\right) \in \mathbb{R}^{d} .
$$

The subscript ${ }_{\text {euc }}$ may be omitted if it is clear from the context that the euclidean inner product is meant. An inner product, and in particular the euclidean inner product on $\mathbb{R}^{d}$, induces a norm by setting $\|u\|:=\langle u, u\rangle^{1 / 2}$.

We denote by $\left(\mathbb{R}^{d}\right)^{\prime}$ the dual space of $\mathbb{R}^{d}$, that is, the space of all linear functionals $\mathbb{R}^{d} \rightarrow \mathbb{R}$. Given a linear functional $u^{\prime} \in\left(\mathbb{R}^{d}\right)^{\prime}$, we write

$$
u^{\prime}(u) \text { or } u^{\prime} u \text { or }\left\langle u^{\prime}, u\right\rangle \text { or }\left\langle u^{\prime}, u\right\rangle_{\left(\mathbb{R}^{d}\right)^{\prime}, \mathbb{R}^{d}}
$$

for the value of $u^{\prime}$ at the point $u$. The dual space $\left(\mathbb{R}^{d}\right)^{\prime}$ is equipped with the norm

$$
\left\|u^{\prime}\right\|_{\left(\mathbb{R}^{d}\right)^{\prime}}=\sup _{\|u\|_{\mathbb{R}^{d}} \leq 1}\left|\left\langle u^{\prime}, u\right\rangle\right| .
$$

It may happen that we simply use $\|\cdot\|$ for the norm both on $\mathbb{R}^{d}$ and on $\left(\mathbb{R}^{d}\right)^{\prime}$, for example, if it is clear from the context, and especially from the element to which the norm is applied, which norm is meant. Similarly, there will be no confusion if we use $\langle\cdot, \cdot\rangle$ for an inner product of two elements of $\mathbb{R}^{d}$, and for the duality between an element in $\left(\mathbb{R}^{d}\right)^{\prime}$ and an element in $\mathbb{R}^{d}$.

Continuity of functions from or to $\mathbb{R}^{d}$ or $\left(\mathbb{R}^{d}\right)^{\prime}$ is always meant with respect to the topology induced by a norm. Recall on this occasion that on a given finite dimensional space like $\mathbb{R}^{d}$ or $\left(\mathbb{R}^{d}\right)^{\prime}$ there is in fact only one such topology since any two norms on this space are equivalent.

Lemma 1.1 (Representation lemma). For every linear functional $u^{\prime} \in\left(\mathbb{R}^{d}\right)^{\prime}$ there exists a unique vector $u \in \mathbb{R}^{d}$ such that

$$
\begin{equation*}
u^{\prime}(v)=\langle u, v\rangle_{\text {euc }} \quad \text { for every } v \in \mathbb{R}^{d} \tag{1.1}
\end{equation*}
$$

Proof. For every $u \in \mathbb{R}^{d}$ we define a linear functional $u^{\prime} \in\left(\mathbb{R}^{d}\right)^{\prime}$ by

$$
u^{\prime}(v):=\langle u, v\rangle_{e u c}, \quad v \in \mathbb{R}^{d} .
$$

Due to the bilinearity and definiteness of the euclidean inner product, the mapping

$$
\begin{aligned}
J: \mathbb{R}^{d} & \rightarrow\left(\mathbb{R}^{d}\right)^{\prime} \\
u & \mapsto u^{\prime}
\end{aligned}
$$

thus defined is linear and injective. Since $\mathbb{R}^{d}$ and $\left(\mathbb{R}^{d}\right)^{\prime}$ are both $d$-dimensional $(d<$ $\infty$ ), the mapping $J$ is actually a linear isomorphism. This means that every linear functional $u^{\prime}$ on $\mathbb{R}^{d}$ is represented by some unique element $u \in \mathbb{R}^{d}$ via the identity (1.1) and the lemma is proved.

### 1.2 Fréchet derivative

Let $U$ be an open subset of $\mathbb{R}^{d}$, and let $\mathscr{E}: U \rightarrow \mathbb{R}$ be a function. We say that $\mathscr{E}$ is differentiable if for every $u \in U$ there exists a linear functional $u^{\prime} \in\left(\mathbb{R}^{d}\right)^{\prime}$ such that

$$
\begin{equation*}
\lim _{\|h\| \rightarrow 0} \frac{\mathscr{E}(u+h)-\mathscr{E}(u)-\left\langle u^{\prime}, h\right\rangle}{\|h\|}=0 . \tag{1.2}
\end{equation*}
$$

By definition, if a function is differentiable, then in the neighbourhood of every point $u \in U$ it can be written as the sum of a constant term, a linear term, and a rest of order $o(h)$. The function $h \mapsto \mathscr{E}(u+h)$ is approximated by the affine function $h \mapsto \mathscr{E}(u)+\left\langle u^{\prime}, h\right\rangle$, with a rest of order $o(h)$, so that a particular case of Taylor's theorem (Taylor expansion up to order 1) is true. The functional $u^{\prime}$, which represents the linear term, is uniquely determined (exercise!).

The derivative is the function function $\mathscr{E}^{\prime}: U \rightarrow\left(\mathbb{R}^{d}\right)^{\prime}$ which assigns to every $u \in U$ the unique linear functional $u^{\prime} \in\left(\mathbb{R}^{d}\right)^{\prime}$ for which (1.2) holds. Consequently, we denote the derivative of $\mathscr{E}$ at a point $u$ by $\mathscr{E}^{\prime}(u)$. We say that $\mathscr{E}$ is continuously differentiable if $\mathscr{E}$ is differentiable and if the derivative is continuous from $U$ into $\left(\mathbb{R}^{d}\right)^{\prime}$. The set of all continuously differentiable functions $U \rightarrow \mathbb{R}$ is a vector space which is denoted by $C^{1}(U)$.

Given $u \in U$, we say that the function $\mathscr{E}$ admits a directional derivative in direction $h \in \mathbb{R}^{d}$ if the limit

$$
\frac{\partial \mathscr{E}}{\partial h}(u):=\lim _{t \rightarrow 0} \frac{\mathscr{E}(u+t h)-\mathscr{E}(u)}{t}
$$

exists. If $h=e_{i}$ is a canonical basis vector, then we call $\frac{\partial \mathscr{E}}{\partial e_{i}}(u)=: \frac{\partial \mathscr{E}}{\partial u_{i}}(u)$ also the partial derivative of $\mathscr{E}$ with respect to $e_{i}$. It is an exercise to show that if $\mathscr{E}$ is differentiable, then $\mathscr{E}$ admits at every $u \in U$ directional derivatives in all directions $h \in \mathbb{R}^{d}$ and

$$
\frac{\partial \mathscr{E}}{\partial h}(u)=\mathscr{E}^{\prime}(u) h
$$

### 1.3 Euclidean gradient

There are several ways to introduce the gradient of a differentiable function, and the most familiar but least flexible is perhaps the one from Lemma 1.2 below. But the gradient is an object, as we already pointed out, which depends on the geometry of the ambient space. This is euclidean space here, with the euclidean inner product and the euclidean norm. The following definition seems therefore to be more appropriate.

Let $U$ be an open subset of $\mathbb{R}^{d}$, and let $\mathscr{E}: U \rightarrow \mathbb{R}$ be a differentiable function. The euclidean gradient of $\mathscr{E}$ is the function $\nabla_{e u c} \mathscr{E}$ which assigns to every point $u \in U$ the unique element $\nabla_{\text {euc }} \mathscr{E}(u) \in \mathbb{R}^{d}$ such that

$$
\begin{equation*}
\mathscr{E}^{\prime}(u) v=\left\langle\nabla_{\text {euc }} \mathscr{E}(u), v\right\rangle_{\text {euc }} \quad \text { for every } v \in \mathbb{R}^{d} \tag{1.3}
\end{equation*}
$$

By the Representation Lemma (Lemma 1.1) the euclidean gradient $\nabla_{\text {euc }} \mathscr{E}$ is well defined in the sense that it exists and that it is unique. We emphasize that the euclidean gradient at some point $u$ is an element of $\mathbb{R}^{d}$ and that this vector has to be distinguished from the derivative $\mathscr{E}^{\prime}(u)$ which is an element of the dual space $\left(\mathbb{R}^{d}\right)^{\prime}$.

It should be clear that the two sets $\mathbb{R}^{d}$ and $\left(\mathbb{R}^{d}\right)^{\prime}$ are different from the set theoretical point of view. On the other hand, as it is often done in mathematics, one may identify the two spaces via some linear isomorphism. One natural isomorphism is given via the euclidean inner product $\langle\cdot, \cdot\rangle_{\text {euc }}$ on $\mathbb{R}^{d}$ and the Representation Lemma (Lemma 1.1).

The following lemma gives an alternative expression of the euclidean gradient in terms of partial derivatives.

Lemma 1.2 (Euclidean gradient and partial derivatives). For every $u \in U$,

$$
\nabla_{e u c} \mathscr{E}(u)=\left(\frac{\partial \mathscr{E}}{\partial e_{1}}(u), \ldots, \frac{\partial \mathscr{E}}{\partial e_{d}}(u)\right) .
$$

Proof. Let $\left(\nabla_{\text {euc }} \mathscr{E}(u)\right)_{i}$ be the $i$-th component of the euclidean gradient of $\mathscr{E}$. By definition of the euclidean inner product and by definition of the canonical basis vectors $e_{i}$ we have $\left(\nabla_{\text {euc }} \mathscr{E}(u)\right)_{i}=\left\langle\nabla_{\text {euc }} \mathscr{E}(u), e_{i}\right\rangle_{\text {euc }}$. Moreover, by definition of the partial derivatives and the euclidean gradient, for every $1 \leq i \leq d$,

$$
\begin{aligned}
\frac{\partial \mathscr{E}}{\partial e_{i}}(u) & =\mathscr{E}^{\prime}(u) e_{i} \\
& =\left\langle\nabla_{\text {euc }} \mathscr{E}(u), e_{i}\right\rangle_{e u c} \\
& =\left(\nabla_{\text {euc }} \mathscr{E}(u)\right)_{i}
\end{aligned}
$$

This was the claim.
More important than the previous identification of the gradient is the property that the gradient direction coincides with the direction of steepest ascent, that is the direction into which $\mathscr{E}$ has greatest directional derivative. The last line of the
following lemma characterizes the euclidean gradient in terms of the derivative and the steepest ascent direction.

Lemma 1.3 (Euclidean gradient and steepest ascent). Assume that $\mathscr{E}$ is differentiable at $u \in U$ and that $\mathscr{E}^{\prime}(u) \neq 0$. Then there exists a unique steepest ascent direction, that is, there exists a unique vector $v \in \mathbb{R}^{d}$ with $\|v\|=1$ such that

$$
\left\|\mathscr{E}^{\prime}(u)\right\|=\sup _{\|w\|=1} \mathscr{E}^{\prime}(u) w=\mathscr{E}^{\prime}(u) v
$$

Given this steepest ascent direction v one has

$$
\nabla_{e u c} \mathscr{E}(u)=\left\|\mathscr{E}^{\prime}(u)\right\| v
$$

Proof. This statement is a consequence of the Cauchy-Schwarz inequality. By definition of the gradient and by the Cauchy-Schwarz inequality we have for $\|w\|=1$

$$
\mathscr{E}^{\prime}(u) w=\left\langle\nabla_{\text {euc }} \mathscr{E}(u), w\right\rangle_{\text {euc }} \leq\left\|\nabla_{\text {euc }} \mathscr{E}(u)\right\|\|w\|=\left\|\nabla_{\text {euc }} \mathscr{E}(u)\right\|,
$$

and equality holds if and only if $\nabla_{e u c} \mathscr{E}(u)$ and $w$ are colinear. Hence, if $\mathscr{E}^{\prime}(u) \neq 0$, then $v=\nabla_{\text {euc }} \mathscr{E}(u) /\left\|\nabla_{\text {euc }} \mathscr{E}(u)\right\|$ is the desired unique vector in the unit ball which maximizes the directional derivatives of $\mathscr{E}$.

Lemma 1.4. Let $U \subseteq \mathbb{R}^{d}$ be an open set, and let $\mathscr{E}: U \rightarrow \mathbb{R}$ be a differentiable function. Then $\mathscr{E}$ is continuously differentiable if and only if the euclidean gradient $\nabla_{\text {euc }} \mathscr{E}: U \rightarrow \mathbb{R}^{d}$ is continuous.

Proof. Let $J: \mathbb{R}^{d} \rightarrow\left(\mathbb{R}^{d}\right)^{\prime}$ be the isomorphism from the proof of the Representation Lemma (Lemma 1.1). With this isomorphism one has $\mathscr{E}^{\prime}=J \nabla_{\text {euc }} \mathscr{E}$, or $\nabla_{\text {euc }} \mathscr{E}=$ $J^{-1} \mathscr{E}^{\prime}$. The claim follows from these two equalities and the fact that $J$ and $J^{-1}$ are continuous.

### 1.4 Euclidean gradient systems

A euclidean gradient system is an ordinary differential equation of the form

$$
\begin{equation*}
\dot{u}+\nabla_{\text {euc }} \mathscr{E}(u)=0 \tag{1.4}
\end{equation*}
$$

where $\mathscr{E}$ is a given continuously differentiable, real-valued function on an open subset $U$ of $\mathbb{R}^{d}$. The unknown in this equation is an $\mathbb{R}^{d}$-valued function $u$ defined on an interval $I \subseteq \mathbb{R}$, which is in general also unknown.

A euclidean gradient system is a special case of the general ordinary differential equation

$$
\begin{equation*}
\dot{u}+F(t, u)=0, \tag{1.5}
\end{equation*}
$$

where $F: D \rightarrow \mathbb{R}^{d}$ is a function on an open set $D \subseteq \mathbb{R}^{1+d}$; consider $F(t, u)=\nabla_{\text {euc }} \mathscr{E}(u)$ for $(t, u) \in D=\mathbb{R} \times U$. In the general ordinary differential equation, the unknown is again a function $u: I \rightarrow \mathbb{R}^{d}$ defined on an interval $I \subseteq \mathbb{R}$. The equalities (1.4) and (1.5) have to be understood as equalities of functions defined on some interval $I$. In this way, we avoid to write the argument of $u$ in both differential equations. The argument of $u$ will in the following usually be denoted by $t$ and is interpreted as time, although in concrete applications it may have a different interpretation. Together with the argument, the gradient system has the form $\dot{u}(t)+\nabla_{e u c} \mathscr{E}(u(t))=0$.

We call a function $u: I \rightarrow \mathbb{R}^{d}$ a solution of the differential equation (1.5) if $u$ is continuous, $F(\cdot, u(\cdot))$ is locally integrable, and if

$$
u(t)=u(s)-\int_{s}^{t} F(r, u(r)) d r \quad \text { for every } s, t \in I
$$

Since the gradient system (1.4) is a special case of the ordinary differential equation (1.5), a solution of the gradient system (1.4) is therefore a continuous function $u$ : $I \rightarrow \mathbb{R}^{d}$ such that

$$
u(t)=u(s)-\int_{s}^{t} \nabla_{e u c} \mathscr{E}(u(r)) d r \quad \text { for every } s, t \in I
$$

A solution in the above sense is a priori not differentiable in the classical sense and is not a solution in the sense which one might expect from the formulation of the equation. Instead, a solution satisfies the differential equations (1.4) and (1.5) only in an integrated form. However, if $\mathscr{E}$ is continuously differentiable as we assumed, then its euclidean gradient is continuous by Lemma 1.4, and therefore, by the preceding equality and the fundamental theorem of calculus, a solution of the gradient system is necessarily continuously differentiable and satisfies the equation $\dot{u}+\nabla_{e u c} \mathscr{E}(u)=0$ pointwise everywhere. However, if $u$ is a solution of the differential equation $\dot{u}+F(t, u)=0$ and if the function $F$ satisfies only weak regularity conditions ( $F$ is not necessarily continuous), then there may exist solutions which are not differentiable in the classical sense.

If $u$ is a solution of the gradient system (1.4), then the time derivative of $u$ is always equal to the negative gradient $-\nabla \mathscr{E}(u)$ and therefore points into the direction of steepest descent (Lemma 1.3). It is therefore natural to expect that the energy $\mathscr{E}$ is decreasing along any solution of the gradient system. The proof of this fact has been given in the Introduction and we recall it here.

Lemma 1.5. Whenever $u$ is a solution of the gradient system (1.4), and $\mathscr{E}$ is continuously differentiable, then the composition $\mathscr{E}(u)=\mathscr{E} \circ u$ is a decreasing function. Moreover, if the composition $\mathscr{E}(u)$ is constant, then the solution $u$ itself is constant.

Proof. Since the function $\mathscr{E}$ and the solution $u$ are continuously differentiable, the composition $\mathscr{E}(u)$ is continuously differentiable, too. Hence, it suffices to prove that
the derivative of $\mathscr{E}(u)$ is nonpositive. For this, we simply calculate the derivative of this function:

$$
\begin{aligned}
\frac{d}{d t} \mathscr{E}(u) & =\mathscr{E}^{\prime}(u) \dot{u} & & \text { (chain rule) } \\
& =\left\langle\nabla_{e u c} \mathscr{E}(u), \dot{u}\right\rangle_{e u c} & & \text { (definition of the euclidean gradient) } \\
& =-\langle\dot{u}, \dot{u}\rangle_{e u c} & & \text { (u is a solution of the gradient system) } \\
& \leq 0 & & \text { (positivity of the euclidean inner product). }
\end{aligned}
$$

This inequality implies that $\mathscr{E}(u)$ is decreasing, but it also implies that if $\mathscr{E}(u)$ is constant, then $\langle\dot{u}, \dot{u}\rangle_{\text {euc }}=0$. Since the euclidean inner product is definite, one then obtains $\dot{u}=0$, and therefore $u$ is constant, too.

Let $U \subseteq \mathbb{R}^{d}$ be an open set and let $F: U \rightarrow \mathbb{R}^{d}$ be a continuous function. We call a function $\mathscr{E}: U \rightarrow \mathbb{R}$ an energy function for the ordinary differential equation

$$
\begin{equation*}
\dot{u}+F(u)=0 \tag{1.6}
\end{equation*}
$$

if for every solution $u$ of this differential equation the composition $\mathscr{E}(u)=\mathscr{E} \circ u$ is decreasing. By the preceding lemma, $\mathscr{E}$ is always an energy function for the gradient system (1.4). The existence of an energy function is quite common for ordinary differential equations coming from applications. For example, in ordinary differential equations arising from physical models it is natural that some physical energy is decreasing or conserved along solutions. These models are certainly a motivation for the term "energy function". Very often, an energy function is also called Lyapunov function. In some places of the literature, an energy function is also called cost function, mainly because of applications to optimisation problems. Differential equations admitting an energy function may be called dissipative systems ${ }^{1}$.

Lemma 1.6. Every limit point of a global solution of the euclidean gradient system (1.4) is an equilibrium point of $\mathscr{E}$.

In other words: if $u: \mathbb{R}_{+} \rightarrow \mathbb{R}^{d}$ is a solution, if $\varphi=\lim _{n \rightarrow \infty} u\left(t_{n}\right)$ exists for some $\left(t_{n}\right) \nearrow \infty$, and if $\varphi \in U$, then $\nabla_{\text {euc }} \mathscr{E}(\varphi)=0$.

Proof. By the preceding lemma, the composed function $\mathscr{E}(u)$ is decreasing. Moreover, since $\lim _{n \rightarrow \infty} \mathscr{E}\left(u\left(t_{n}\right)\right)=\mathscr{E}(\varphi)$ exists by assumption and by continuity of $\mathscr{E}$, it follows that $\mathscr{E}(u)$ is bounded below. Integrating the equality $\frac{d}{d t} \mathscr{E}(u)=-\|\dot{u}\|^{2}$ we then obtain that the integral $\int_{0}^{\infty}\|\dot{u}\|^{2}$ is finite. This implies $\lim _{t \rightarrow \infty} \int_{t}^{t+1}\|\dot{u}\|^{2}=0$.

Hence,

[^2]$$
\lim _{n \rightarrow \infty} u\left(t_{n}+s\right)=\lim _{n \rightarrow \infty}\left(u\left(t_{n}\right)+\int_{t_{n}}^{t_{n}+s} \dot{u}\right)=\varphi \quad \text { uniformly in } s \in[0,1]
$$
since, by Hölder's inequality,
$$
\sup _{s \in[0,1]}\left\|\int_{t_{n}}^{t_{n}+s} \dot{u}\right\| \leq \sup _{s \in[0,1]} \int_{t_{n}}^{t_{n}+s}\|\dot{u}\| \leq\left(\int_{t_{n}}^{t_{n}+1}\|\dot{u}\|^{2}\right)^{\frac{1}{2}} \rightarrow 0 \quad \text { as } n \rightarrow \infty .
$$

By continuity of $\nabla_{\text {euc }} \mathscr{E}$, this implies $\lim _{n \rightarrow \infty} \nabla_{e u \mathscr{C}} \mathscr{E}\left(u\left(t_{n}+s\right)\right)=\nabla_{e u c} \mathscr{E}(\varphi)$ uniformly in $s \in[0,1]$, and therefore

$$
\begin{aligned}
\nabla_{e u c} \mathscr{E}(\varphi) & =\int_{0}^{1} \nabla_{\text {euc }} \mathscr{E}(\varphi) \\
& =\lim _{n \rightarrow \infty} \int_{0}^{1} \nabla_{\text {euc }} \mathscr{E}\left(u\left(t_{n}+s\right)\right) d s \\
& =-\lim _{n \rightarrow \infty} \int_{0}^{1} \dot{u}\left(t_{n}+s\right) d s \\
& =0
\end{aligned}
$$

This is the claim.

### 1.5 Algebraic equations and steepest descent

Let $U \subseteq \mathbb{R}^{d}$ be an open set and let $F: U \rightarrow \mathbb{R}^{d}$ be a continuously differentiable function. Consider the problem of finding a solution $\bar{u} \in U$ of the algebraic equation

$$
F(\bar{u})=0
$$

One way to solve this problem could be to consider the function $\mathscr{E}: U \rightarrow \mathbb{R}$ given by $\mathscr{E}(u)=\frac{1}{2}\|F(u)\|^{2}$, where $\|\cdot\|$ is the euclidean norm. Then one has $F(\bar{u})=0$ if and only if $\mathscr{E}(\bar{u})=0$. Since $\mathscr{E}$ is a positive function, the problem amounts to search for a point where the infimum of $\mathscr{E}$ is attained and, if it is attained, to hope that this infimum is equal to 0 . The infimum of $\mathscr{E}$ can be searched for by a gradient descent method $^{2}$.

It is an exercice to show that the derivative of the function $\mathscr{E}$ is given by

$$
\mathscr{E}^{\prime}(u) v=\left\langle F(u), F^{\prime}(u) v\right\rangle \quad \text { for every } u \in U, v \in \mathbb{R}^{d}
$$

We define the adjoint of $F^{\prime}(u)$ with respect to the euclidean inner product to be the linear mapping $F^{\prime}(u)^{*}: \mathbb{R}^{d} \rightarrow \mathbb{R}^{d}$ given by

[^3]$$
\left\langle F^{\prime}(u)^{*} v, w\right\rangle=\left\langle v, F^{\prime}(u) w\right\rangle \quad \text { for every } v, w \in \mathbb{R}^{d}
$$

With this definition, the euclidean gradient of the function $\mathscr{E}$ is given by

$$
\nabla \mathscr{E}(u)=F^{\prime}(u)^{*} F(u),
$$

and the euclidean gradient system associated with $\mathscr{E}$ is hence the differential equation

$$
\begin{equation*}
\dot{u}+F^{\prime}(u)^{*} F(u)=0 . \tag{1.7}
\end{equation*}
$$

Proposition 1.7. Assume that $F^{\prime}(v)$ is invertible for every $v \in U$. Assume further that there exists a (global) solution $u: \mathbb{R}_{+} \rightarrow \mathbb{R}^{d}$ of (1.7) with relatively compact range in $U$. Then there exists $\bar{u} \in U$ such that $F(\bar{u})=0$.

Proof. Since the solution $u$ has relatively compact range in $U$, there exist a sequence $\left(t_{n}\right) \nearrow \infty$ and an element $\bar{u} \in U$ such that $\lim _{n \rightarrow \infty} u\left(t_{n}\right)=\bar{u}$. By what has been shown above, the differential equation (1.7) is the euclidean gradient system associated with the function $\mathscr{E}(u)=\frac{1}{2}\|F(u)\|^{2}$. By Lemma 1.6, the element $\bar{u} \in U$ is therefore a critical point of $\mathscr{E}$, that is $F^{\prime}(\bar{u})^{*} F(\bar{u})=0$. Since $F^{\prime}(\bar{u})$, and therefore also $F^{\prime}(\bar{u})^{*}$, is invertible by assumption, this implies $F(\bar{u})=0$.

From the proof of the preceding proposition it turns out that every limit point of every global solution of the gradient system (1.7) having relatively compact range in $U$ is a solution to the algebraic equation $F(\bar{u})=0$. In other words, the problem of finding a solution to an algebraic equation is transformed into the problem of finding a global solution of a gradient system having relatively compact range.

We will study the problem of global existence for the first time in the following lecture and we will see that the gradient structure also helps for that problem.

### 1.6 Exercises

1.1. The sublevels of an energy function $\mathscr{E}: \mathbb{R}^{d} \rightarrow \mathbb{R}$ are the sets $K_{c}=\left\{u \in \mathbb{R}^{d}\right.$ : $\mathscr{E}(u) \leq c\}$ with $c \in \mathbb{R}$. Show that each sublevel is invariant under the gradient system $\dot{u}+\nabla_{\text {euc }} \mathscr{E}(u)=0$, that is, if $u:[0, T] \rightarrow \mathbb{R}^{d}$ is a solution and $u(0) \in K_{c}$, then $u(t) \in K_{c}$ for every $t \in[0, T]$.
1.2. a) Prove that every ordinary differential equation of the form $\dot{u}+F(u)=0$ with continuous $F: \mathbb{R} \rightarrow \mathbb{R}$ is a euclidean gradient system.
b) Prove that every solution of the above ordinary differential equation is monotone.
1.3. a) Prove that the system

$$
\begin{aligned}
\dot{u}_{1}+u_{1}^{2} u_{2} & =0 \\
\dot{u}_{2}+\frac{1}{3} u_{1}^{3} & =0
\end{aligned}
$$

is a euclidean gradient system.
b) Prove that the differential system of the damped pendulum

$$
\begin{array}{r}
\dot{u}_{1}-u_{2}=0 \\
\dot{u}_{2}+\alpha u_{2}+\beta \sin u_{1}=0
\end{array}
$$

$\alpha>0, \beta>0$, is not a euclidean gradient system.
1.4. Let $\mathbb{R}^{d}$ be equipped with a norm which is strictly convex, that is, for any two points $x, y \in \mathbb{R}^{d}$ with $\|x\|=\|y\|=1$ and $x \neq y$ one has $\left\|\frac{x+y}{2}\right\|<1$. Let $\mathscr{E}: \mathbb{R}^{d} \rightarrow \mathbb{R}$ be a differentiable function. Prove that for every $u \in \mathbb{R}^{d}$ there exists a unique vector $v \in \mathbb{R}^{d},\|v\|=1$, such that

$$
\left\|\mathscr{E}^{\prime}(u)\right\|:=\sup _{\|w\|=1} \mathscr{E}^{\prime}(u) w=\mathscr{E}^{\prime}(u) v
$$

Given this vector $v$, we define the gradient $\nabla \mathscr{E}(u):=\left\|\mathscr{E}^{\prime}(u)\right\| v$. By Lemma 1.3, this gradient coincides with the euclidean gradient if $\|\cdot\|=\|\cdot\|_{e u c}$. Prove that the euclidean norm is strictly convex.

Remark: The $p$-norm given by $\|u\|_{p}=\left(\sum_{i}\left|u_{i}\right|^{p}\right)^{\frac{1}{p}}$ is strictly convex if $1<p<\infty$.
1.5. Consider the differential equation

$$
\dot{u}+\frac{\nabla \mathscr{E}(u)}{\|\nabla \mathscr{E}(u)\|}=0
$$

and show that solutions of this differential equation and solutions of the gradient system $\dot{v}+\nabla \mathscr{E}(v)=0$ can be transformed into each other by a change of time variable $u(t)=v(\alpha(t))$ and $v(t)=u(\beta(t))$ for some functions $\alpha$ and $\beta$ depending on $v$ and $u$, respectively. Find the differential equations determining the functions $\alpha$ and $\beta$.
1.6. Draw the phase portrait (or alternatively: the direction field) of some planar gradient system and insert the level curves of the underlying energy function. What do you observe?

## Lecture 2

## Gradient systems in finite dimensional space

In this lecture we continue to consider functions, gradients and gradient systems on $\mathbb{R}^{d}$. But we generalize the notion of gradient and gradient system in two steps. The key is to replace the euclidean inner product in the definition of the euclidean gradient by an arbitrary inner product or by a Riemannian metric.

Recall that an inner product on $\mathbb{R}^{d}$ is a bilinear, symmetric, positive definite form $\langle\cdot, \cdot\rangle: \mathbb{R}^{d} \times \mathbb{R}^{d} \rightarrow \mathbb{R}$, that is, for every $u, v, w \in \mathbb{R}^{d}$ and every $\lambda \in \mathbb{R}$ one has
(i) $\langle\lambda u+v, w\rangle=\lambda\langle u, w\rangle+\langle v, w\rangle$ and $\langle u, \lambda v+w\rangle=\lambda\langle u, v\rangle+\langle u, w\rangle$ (bilinearity),
(ii) $\langle u, v\rangle=\langle v, u\rangle$ (symmetry) and
(iii) $\langle u, u\rangle \geq 0$ (positivity) with equality only for $u=0$ (definiteness).

In the following, we denote by $\mathscr{L}_{2}\left(\mathbb{R}^{d} ; \mathbb{R}\right)$ the space of all bilinear forms $a: \mathbb{R}^{d} \times$ $\mathbb{R}^{d} \rightarrow \mathbb{R}$, that is, of all forms satisfying property (i) above. This space is a vector space for the natural addition and the natural scalar multiplication, and it is a Banach space for the norm

$$
\|a\|:=\sup _{\substack{\|u\| \leq 1 \\\|v\| \leq 1}}|a(u, v)| .
$$

The set of all inner products on $\mathbb{R}^{d}$, which we denote by $\operatorname{Inner}\left(\mathbb{R}^{d}\right)$, is a subset of $\mathscr{L}_{2}\left(\mathbb{R}^{d} ; \mathbb{R}\right)$. It is therefore naturally equipped with the metric induced by the above norm.

Lemma 2.1 (Representation lemma). Let $\langle\cdot, \cdot\rangle$ be an inner product on $\mathbb{R}^{d}$. For every linear functional $u^{\prime} \in\left(\mathbb{R}^{d}\right)^{\prime}$ there exists a unique vector $u \in \mathbb{R}^{d}$ such that

$$
\begin{equation*}
u^{\prime}(v)=\langle u, v\rangle \quad \text { for every } v \in \mathbb{R}^{d} \tag{2.1}
\end{equation*}
$$

We say that the functional $u^{\prime} \in\left(\mathbb{R}^{d}\right)^{\prime}$ is represented by the element $u \in \mathbb{R}^{d}$ with respect to the inner product $\langle\cdot, \cdot\rangle$.

Proof. Note that for every $u \in \mathbb{R}^{d}$ one may define a linear functional $u^{\prime} \in\left(\mathbb{R}^{d}\right)^{\prime}$ by

$$
u^{\prime}(v):=\langle u, v\rangle, \quad v \in \mathbb{R}^{d} .
$$

Due to the bilinearity and definiteness of the inner product $\langle\cdot, \cdot\rangle$, the mapping

$$
\begin{aligned}
J: \mathbb{R}^{d} & \rightarrow\left(\mathbb{R}^{d}\right)^{\prime} \\
u & \mapsto u^{\prime}
\end{aligned}
$$

thus defined is linear and injective. Since $\mathbb{R}^{d}$ and $\left(\mathbb{R}^{d}\right)^{\prime}$ are both $d$-dimensional $(d<$ $\infty$ ), the mapping $J$ is actually a linear isomorphism. This means that every linear functional $u^{\prime}$ on $\mathbb{R}^{d}$ is represented by some unique element $u \in \mathbb{R}^{d}$ via the identity (2.1) and the lemma is proved.

Lemma 2.2. Let $\langle\cdot, \cdot\rangle$ be an inner product on $\mathbb{R}^{d}$. Then there exists a linear mapping $Q: \mathbb{R}^{d} \rightarrow \mathbb{R}^{d}$ which is
(i) symmetric, that is, for every $v, w \in \mathbb{R}^{d}$ one has $\langle Q v, w\rangle_{e u c}=\langle v, Q w\rangle_{\text {euc }}$,
(ii) positive, that is, for every $v \in \mathbb{R}^{d}$ one has $\langle Q v, v\rangle_{\text {euc }} \geq 0$,
(iii) definite, that is, if $\langle Q v, w\rangle_{\text {euc }}=0$ for every $w \in \mathbb{R}^{d}$, then $v=0$, and
(iv) for every $v, w \in \mathbb{R}^{d}$ one has

$$
\langle v, w\rangle=\langle Q v, w\rangle_{e u c}
$$

Proof. For every $v \in \mathbb{R}^{d}$ the mapping $v^{\prime}: \mathbb{R}^{d} \rightarrow \mathbb{R}$ given by $v^{\prime}(w)=\langle v, w\rangle\left(w \in \mathbb{R}^{d}\right)$ is linear and continuous. The Representation Lemma 2.1 applied to the euclidean inner product (or: the Representation Lemma 1.1) yields that there exists an element $Q v \in \mathbb{R}^{d}$ such that $v^{\prime}(w)=\langle Q v, w\rangle_{e u c}$ for every $w \in \mathbb{R}^{d}$. By definition, $Q$ satisfies property (iv). Moreover, as a consequence of the bilinearity of $\langle\cdot, \cdot\rangle$, the mapping $Q$ is linear. By definition of $Q$ and by the symmetry of the inner products $\langle\cdot, \cdot\rangle$ and $\langle\cdot, \cdot\rangle_{e u c}$,

$$
\langle Q v, w\rangle_{e u c}=\langle v, w\rangle=\langle w, v\rangle=\langle Q w, v\rangle_{e u c}=\langle v, Q w\rangle_{e u c} \quad \text { for every } v, w \in \mathbb{R}^{d}
$$

Hence, $Q$ is symmetric. Moreover, for every $v \in \mathbb{R}^{d}$,

$$
\langle Q v, v\rangle_{e u c}=\langle v, v\rangle \geq 0
$$

by positivity of the inner product $\langle\cdot, \cdot\rangle$. Finally, by definiteness of the inner product $\langle\cdot, \cdot\rangle,\langle Q v, w\rangle_{\text {euc }}=0$ for every $w \in \mathbb{R}^{d}$ implies $v=0$, and hence $Q$ is definite, too.

Let $U \subseteq \mathbb{R}^{d}$ be an open set. A continuous function $g: U \rightarrow \operatorname{Inner}\left(\mathbb{R}^{d}\right)$ is called a Riemannian metric on $U$. Given a Riemannian metric, we write $\langle\cdot, \cdot\rangle_{g(u)}$ to denote the inner product $g(u)$ at the point $u \in U$, and we denote by $\|\cdot\|_{g(u)}$ the corresponding norm on $\mathbb{R}^{d}$.

Lemma 2.3. Let $g: U \rightarrow \operatorname{Inner}\left(\mathbb{R}^{d}\right)$ be a Riemannian metric, and let $Q: U \rightarrow$ $\mathscr{L}\left(\mathbb{R}^{d}\right)$ be the function given by

$$
\langle Q(u) v, w\rangle_{e u c}=\langle v, w\rangle_{g(u)} \quad \text { for every } u \in U, v, w \in \mathbb{R}^{d}
$$

Then $Q$ is continuous.
Proof. For every $u_{1}, u_{2} \in U$ we have

$$
\begin{aligned}
\left\|Q\left(u_{1}\right)-Q\left(u_{2}\right)\right\| & =\sup _{\|v\| \leq 1}\left\|Q\left(u_{1}\right) v-Q\left(u_{2}\right) v\right\| \\
& =\sup _{\substack{\|v \mid \leq 1\\
\| w \| \leq 1}}\left|\left\langle Q\left(u_{1}\right) v-Q\left(u_{2}\right) v, w\right\rangle_{e u c}\right| \\
& =\sup _{\substack{\|v\| \leq 1 \\
\|w\| \leq 1}}\left|\langle v, w\rangle_{g\left(u_{1}\right)}-\langle v, w\rangle_{g\left(u_{2}\right)}\right| \\
& =\left\|g\left(u_{1}\right)-g\left(u_{2}\right)\right\| .
\end{aligned}
$$

Since $g$ is continuous, the continuity of $Q$ follows.

### 2.1 Definition of gradient

Let $U \subseteq \mathbb{R}^{d}$ be an open set and let $\mathscr{E}: U \rightarrow \mathbb{R}$ be a continuously differentiable function. We recall from the preceding chapter that for every $u \in U$ and every $v \in \mathbb{R}^{d}$ the derivative and the euclidean gradient are related by the equality

$$
\mathscr{E}^{\prime}(u) v=\left\langle\nabla_{e u c} \mathscr{E}(u), v\right\rangle_{e u c}
$$

Starting from this definition and from the general Representation Lemma 2.1, it is natural to generalize the notion of gradient. We actually consider two generalizations.

First, given an arbitrary inner product $\langle\cdot, \cdot\rangle$ on $\mathbb{R}^{d}$, the gradient of $\mathscr{E}$ with respect to this inner product is the function $\nabla \mathscr{E}$ which assigns to each $u \in U$ the unique element $\nabla \mathscr{E}(u) \in \mathbb{R}^{d}$ such that

$$
\mathscr{E}^{\prime}(u) v=\langle\nabla \mathscr{E}(u), v\rangle \quad \text { for every } v \in \mathbb{R}^{d}
$$

Second, given a Riemannian metric $g$ on $U$, the gradient of $\mathscr{E}$ with respect to $g$ is the function $\nabla_{g} \mathscr{E}$ which assigns to each $u \in U$ the unique element $\nabla_{g} \mathscr{E}(u) \in \mathbb{R}^{d}$ such that

$$
\mathscr{E}^{\prime}(u) v=\left\langle\nabla_{g} \mathscr{E}(u), v\right\rangle_{g(u)} \quad \text { for every } v \in \mathbb{R}^{d}
$$

If the metric $g$ is clear from the context, then we may simply write $\nabla \mathscr{E}$ instead of $\nabla_{g} \mathscr{E}$. Clearly, the second generalization includes the first one: it suffices to take a constant Riemannian metric.

Lemma 2.4. If $\mathscr{E}: U \rightarrow \mathbb{R}$ is continuously differentiable, and if $g$ is a Riemannian metric on $U$, then the gradient $\nabla_{g} \mathscr{E}: U \rightarrow \mathbb{R}^{d}$ is continuous.

Proof. If $Q$ is the function from Lemma 2.3, then $\nabla_{g} \mathscr{E}(u)=Q(u)^{-1} \nabla_{e u c} \mathscr{E}(u)$. Continuity of $\nabla_{g} \mathscr{E}$ is therefore a direct consequence of Lemma 2.3 and Lemma 1.4.

### 2.2 Definition of gradient system

We call a gradient system any ordinary differential equation in $\mathbb{R}^{d}$ which is of the form

$$
\begin{equation*}
\dot{u}+\nabla_{g} \mathscr{E}(u)=0 \tag{2.2}
\end{equation*}
$$

where $\nabla_{g} \mathscr{E}$ is the gradient of a continuously differentiable function $\mathscr{E}: U \rightarrow \mathbb{R}$ with respect to a Riemannian metric $g$. This gradient system admits $\mathscr{E}$ as an energy function. The proof is a repetition of the proof of Lemma 1.5.

Lemma 2.5. Whenever $u$ is a solution of the gradient system (2.2), and $\mathscr{E}$ is continuously differentiable, then the composition $\mathscr{E}(u)=\mathscr{E} \circ u$ is a decreasing function. Moreover, if the composition $\mathscr{E}(u)$ is constant, then the solution u itself is constant.

Proof. Since the function $\mathscr{E}$ and the solution $u$ are continuously differentiable, the composition $\mathscr{E}(u)$ is continuously differentiable, too. Hence, it suffices to prove that the derivative of $\mathscr{E}(u)$ is non-positive.

Using the chain rule, the definition of the gradient $\nabla_{g} \mathscr{E}$, and the differential equation (2.2), we can calculate

$$
\begin{aligned}
\frac{d}{d t} \mathscr{E}(u) & =\mathscr{E}^{\prime}(u) \dot{u} & & \text { (chain rule) } \\
& =\left\langle\nabla_{g} \mathscr{E}(u), \dot{u}\right\rangle_{g(u)} & & \text { (definition of the gradient) } \\
& =-\langle\dot{u}, \dot{u}\rangle_{g(u)} & & (u \text { is solution of the gradient system) } \\
& \leq 0 & & \text { (positivity of the inner product } \left.\langle\cdot, \cdot\rangle_{g(u)}\right) .
\end{aligned}
$$

This inequality implies that $\mathscr{E}(u)$ is decreasing, but it also implies that if $\mathscr{E}(u)$ is constant, then $\langle\dot{u}, \dot{u}\rangle_{g(u)}=0$. Since $\langle\cdot, \cdot\rangle_{g(u)}$ is definite, one then obtains $\dot{u}=0$, and therefore $u$ is constant, too.

### 2.3 Local existence for ordinary differential equations

This section is a small digression. In view of the existence results for gradient systems in infinite dimensional spaces, we discuss in some detail the question of local existence of solutions of gradient systems in finite dimensions. However, since local existence for general ordinary differential equations is not more difficult to prove, we state the general result. The following classical result from the theory of ordinary differential equations, Carathéodory's theorem, is perhaps known and one may skip the proof in this case.

Theorem 2.6 (Carathéodory). Let $D \subseteq \mathbb{R} \times \mathbb{R}^{d}$ be an open set and let $F: D \rightarrow \mathbb{R}^{d}$, $F=F(t, u)$, be a function which satisfies the following Carathéodory conditions:

$$
\begin{align*}
& F(\cdot, u) \text { is measurable for every } u \text {, }  \tag{2.3}\\
& F(t, \cdot) \text { is continuous for every } t \text {, and }  \tag{2.4}\\
& \text { for every }\left(t_{0}, u_{0}\right) \in D \text { there exist } \alpha>0, r>0 \text { and } m \in L^{1}\left(t_{0}, t_{0}+\alpha\right)  \tag{2.5}\\
& \text { such that }\|F(t, u)\| \leq m(t) \text { for every } t \in\left(t_{0}, t_{0}+\alpha\right), u \in B\left(u_{0}, r\right) \text {. }
\end{align*}
$$

Then, for every $\left(t_{0}, u_{0}\right) \in D$ the ordinary differential equation with initial condition

$$
\left\{\begin{array}{l}
\dot{u}+F(t, u)=0,  \tag{2.6}\\
u\left(t_{0}\right)=u_{0},
\end{array}\right.
$$

admits a local solution. This means that there exists an interval $I=\left[t_{0}, t_{0}+\alpha\right] \subseteq \mathbb{R}$ which is not reduced to one point, and there exists a continuous function $u: I \rightarrow \mathbb{R}^{d}$ such that for every $t \in I$

$$
\begin{equation*}
u(t)=u_{0}-\int_{t_{0}}^{t} F(s, u(s)) d s \tag{2.7}
\end{equation*}
$$

Note that this definition of local solution coincides with the definition from Lecture 1.

We give a "constructive" proof of Carathéodory's theorem in the sense that we first define a sequence of local approximate solutions, which "almost" satisfy the integral equation (2.7), and we then prove that a subsequence of these approximate solutions converges to an exact solution of the differential equation (2.6). The compactness argument behind the existence of such a subsequence is based on the Arzelà-Ascoli theorem.

Proof (Proof of Theorem 2.6). Let $\left(t_{0}, u_{0}\right) \in D$, and let the constants $\alpha>0, r>0$ and the function $m$ be as in the condition (2.5). Define $M(t):=\int_{t_{0}}^{t} m(s) d s$ for $t \in$ $\left[t_{0}, t_{0}+\alpha\right]$. Then the function $M$ is continuous and $M\left(t_{0}\right)=0$. Using this continuity and choosing $\alpha>0$ small enough, we may assume that $|M(t)| \leq r$ for every $t \in$ $\left[t_{0}, t_{0}+\boldsymbol{\alpha}\right]$.

For every $k \geq 2$ we define a function $u_{k} \in C\left(\left[t_{0}, t_{0}+\alpha\right] ; \mathbb{R}^{d}\right)$ by
$u_{k}(t):= \begin{cases}u_{0} & \text { if } t \in\left[t_{0}, t_{0}+\frac{\alpha}{k}\right], \\ u_{0}-\int_{t_{0}}^{t-\frac{\alpha}{k}} F\left(s, u_{k}(s)\right) d s & \text { if } t \in\left(t_{0}+\frac{\alpha}{k}, t_{0}+\frac{2 \alpha}{k}\right], \ldots,\left(t_{0}+\frac{(k-1) \alpha}{k}, t_{0}+\alpha\right] .\end{cases}$
Note first that the functions $u_{k}$ are well defined in the sense that the function $F\left(\cdot, u_{k}(\cdot)\right)$ in the integral is defined. In fact, one proves easily that

$$
\begin{equation*}
\left\|u_{k}(t)-u_{0}\right\| \leq r \quad \text { for every } k \geq 2, t \in\left[t_{0}, t_{0}+\alpha\right] \tag{2.8}
\end{equation*}
$$

Fix $\varepsilon>0$. Since the function $M$ is uniformly continuous on the compact interval $\left[t_{0}, t_{0}+\alpha\right]$, there exists $\delta>0$ such that for every $t, t^{\prime} \in\left[t_{0}, t_{0}+\alpha\right]$ the implication

$$
\left|t-t^{\prime}\right| \leq \delta \quad \Rightarrow \quad\left|M(t)-M\left(t^{\prime}\right)\right| \leq \varepsilon
$$

is true.
Moreover, for every $k \geq 2$ and every $t, t^{\prime} \in\left[t_{0}, t_{0}+\alpha\right], t \geq t^{\prime}$, one has

$$
\begin{aligned}
& \left\|u_{k}(t)-u_{k}\left(t^{\prime}\right)\right\| \leq \\
& \quad \leq \begin{cases}\int_{t^{\prime}-\frac{\alpha}{k}}^{t-\frac{\alpha}{k}}\left\|F\left(s, u_{k}(s)\right)\right\| d s \leq\left|M\left(t-\frac{\alpha}{k}\right)-M\left(t^{\prime}-\frac{\alpha}{k}\right)\right| & \text { if } t \geq t^{\prime} \geq t_{0}+\frac{\alpha}{k} \\
\int_{t_{0}}^{t-\frac{\alpha}{k}}\left\|F\left(s, u_{k}(s)\right)\right\| d s \leq\left|M\left(t-\frac{\alpha}{k}\right)\right| & \text { if } t \geq t_{0}+\frac{\alpha}{k} \geq t^{\prime} \\
0 & \text { if } t_{0}+\frac{\alpha}{k} \geq t \geq t^{\prime}\end{cases}
\end{aligned}
$$

In any case, for every $k \geq 2$ and every $t, t^{\prime} \in\left[t_{0}, t_{0}+\alpha\right]$ with $\left|t-t^{\prime}\right| \leq \delta$ one has

$$
\left\|u_{k}(t)-u_{k}\left(t^{\prime}\right)\right\| \leq \varepsilon
$$

This means that the sequence $\left(u_{k}\right) \in C\left(\left[t_{0}, t_{0}+\alpha\right] ; \mathbb{R}^{d}\right)$ is equicontinuous. Moreover, by the estimate (2.8), this sequence is also uniformly bounded.

Hence, by the Arzelà-Ascoli theorem (Theorem B.43), there exists a subsequence $\left(u_{k_{j}}\right)$ of $\left(u_{k}\right)$ and a continuous function $u:\left[t_{0}, t_{0}+\alpha\right] \rightarrow \mathbb{R}^{d}$ such that $u_{k_{j}}$ converges uniformly on $\left[t_{0}, t_{0}+\alpha\right]$ to $u$.

By condition (2.4), for every $t \in\left[t_{0}, t_{0}+\alpha\right]$ one has

$$
\lim _{j \rightarrow \infty} F\left(t, u_{k_{j}}(t)\right)=F(t, u(t)) .
$$

By the estimate (2.8) and by condition (2.5),

$$
\left\|F\left(t, u_{k}(t)\right)\right\| \leq m(t) \quad \text { for every } k \geq 2, t \in\left[t_{0}, t_{0}+\alpha\right]
$$

Hence, by definition of the functions $u_{k}$ and by the dominated convergence theorem, for every $t \in\left[t_{0}, t_{0}+\alpha\right]$

$$
\begin{aligned}
u(t) & =\lim _{j \rightarrow \infty} u_{k_{j}}(t) \\
& =u_{0}-\lim _{j \rightarrow \infty} \int_{t_{0}}^{t-\frac{\alpha}{k_{j}}} F\left(s, u_{k_{j}}(s)\right) d s \\
& =u_{0}-\int_{t_{0}}^{t} F(s, u(s)) d s
\end{aligned}
$$

Hence, $u$ is a local solution.
Theorem 2.6 is like a key which opens the door to study ordinary differential equations, or like a source which provides us with solutions of ordinary differential equations. The question of local existence, under the Carathéodory assumptions,
is completely solved, and we can now study properties of these solutions. Note, however, that without any further assumptions on the function $F$, one can neither expect existence of global solutions (defined on $\left[t_{0}, \infty\right.$ ), for example), nor uniqueness of local solutions.

Example 2.7. a) The function $u(t)=\frac{1}{1-t}, t \in[0,1)$, is a solution of the ordinary differential equation

$$
\left\{\begin{array}{l}
\dot{u}-u^{2}=0, \\
u(0)=1,
\end{array}\right.
$$

and can not be extended to a global solution (defined on $[0, \infty)$ ).
b) The functions $u(t)=0$ and $u(t)=\frac{1}{4} t^{2}, t \in \mathbb{R}_{+}$, are two distinct solutions of the ordinary differential equation

$$
\left\{\begin{array}{l}
\dot{u}-\sqrt{|u|}=0, \\
u(0)=0 .
\end{array}\right.
$$

By Exercise 1.2, both differential equations in (a) and (b) are gradient systems.
A solution $u:\left[t_{0}, t_{0}+\alpha\right) \rightarrow \mathbb{R}^{d}$ of the ordinary differential equation (2.6) is called a maximal solution if it can not be extended to a solution on a strictly larger interval $\left[t_{0}, t_{0}+\beta\right)$. Local existence of solutions of ordinary differential equations always implies the existence of maximal solutions. The proof of this fact is based on the Lemma of Zorn.

Corollary 2.8. Under the assumptions of Theorem 2.6, for every $\left(t_{0}, u_{0}\right) \in D$ there exists a maximal solution of the problem (2.6).

Proof (Idea). Consider the set $\mathscr{S}$ of all pairs $(\alpha, u)$ of positive numbers $\alpha \in(0, \infty]$ ( $\infty$ is included) and of continuous functions $u:\left[t_{0}, t_{0}+\alpha\right) \rightarrow \mathbb{R}^{d}$ which are solutions of the initial value problem (2.6). By Carathéodory's Theorem, this set is nonempty.

On the set $\mathscr{S}$, we consider the following partial ordering: we say that $(\alpha, u) \leq$ $(\beta, v)$, if $\alpha \leq \beta$ with respect to the usual ordering in $(0, \infty]$, and if $u$ is a restriction of the function $v$ to the interval $\left[t_{0}, t_{0}+\alpha\right)$. We leave it as an exercise to show that $\leq$ is really a partial ordering on $\mathscr{S}$.

We also leave it as an exercise to show that every totally ordered subset of $\mathscr{S}$ admits a maximal element. The claim then follows from Zorn's Lemma.

### 2.4 Local and global existence for gradient systems

From Carathéodory's local existence theorem we immediately obtain the following corollary about local existence of solutions of non-autonomous gradient systems.

Corollary 2.9 (Local existence for non-autonomous gradient systems). Let $U \subseteq$ $\mathbb{R}^{d}$ be an open set, let $\mathscr{E}: U \rightarrow \mathbb{R}$ be a continuously differentiable function, and let
$g: U \rightarrow \operatorname{Inner}\left(\mathbb{R}^{d}\right)$ be a Riemannian metric on $U$. Let $I \subseteq \mathbb{R}$ be an open interval, and let $f \in L^{1}\left(I ; \mathbb{R}^{d}\right)$. Then, for every $t_{0} \in I$ and every $u_{0} \in U$ the problem

$$
\left\{\begin{array}{l}
\dot{u}+\nabla_{g} \mathscr{E}(u)=f  \tag{2.9}\\
u\left(t_{0}\right)=u_{0}
\end{array}\right.
$$

admits a local solution. This means that there exists an interval $J=\left[t_{0}, t_{0}+\alpha\right] \subseteq I$ which is not reduced to one point, and there exists a continuous function $u: J \rightarrow \mathbb{R}^{d}$ such that for every $t \in J$

$$
u(t)-u_{0}+\int_{t_{0}}^{t} \nabla_{g} \mathscr{E}(u(s)) d s=\int_{t_{0}}^{t} f(s) d s
$$

Proof. It suffices to check that the assumptions of Carathéodory's theorem are satisfied. We leave this as an exercise.

Maximal solutions of the gradient system (2.9) are defined in the same way as for the general ordinary differential equation (2.6). It follows from Corollary 2.8 that the gradient system (2.9) always admits a maximal solution. For the general ordinary differential equation as well as for the gradient system an important question is to determine the maximal existence interval or to show that maximal solutions are global. By Example 2.7 (a), global existence of solutions does not always hold, neither for ordinary differential equations nor for gradient systems. But under a natural assumption on the energy, global existence of solutions of gradient systems can be proved via a priori estimates.

A function $\mathscr{E}: \mathbb{R}^{d} \rightarrow \mathbb{R}$ is coercive if for every $c \in \mathbb{R}$ the sub-level set

$$
K_{c}:=\left\{u \in \mathbb{R}^{d}: \mathscr{E}(u) \leq c\right\} \text { is bounded. }
$$

Theorem 2.10 (Global existence for gradient systems). Let $\mathscr{E}: \mathbb{R}^{d} \rightarrow \mathbb{R}$ be a continuously differentiable, coercive function, and let $g$ be a Riemannian metric on $\mathbb{R}^{d}$ such that $\|v\|_{g(u)} \leq C\|v\|_{\text {euc }}$ for every $u, v \in \mathbb{R}^{d}$ and some constant $C \geq 0$. Then, for every $u_{0} \in \mathbb{R}^{d}$ and every continuous $f:[0, T] \rightarrow \mathbb{R}^{d}(T>0)$, every maximal solution $u$ of the non-autonomous gradient system

$$
\left\{\begin{array}{l}
\dot{u}+\nabla_{g} \mathscr{E}(u)=f  \tag{2.10}\\
u(0)=u_{0}
\end{array}\right.
$$

exists on $[0, T]$. Moreover,

$$
\{u(t): t \in[0, T]\} \subseteq K_{c},
$$

where $c=\mathscr{E}\left(u_{0}\right)+\frac{C}{2} \int_{0}^{T}\|f\|_{\text {euc }}^{2}$ depends only on the data $u_{0}, f$ and the embedding constant $C$.

Proof. Let $u:\left[0, T_{\max }\right) \rightarrow \mathbb{R}^{d}\left(T_{\max } \leq T\right)$ be any maximal solution of the nonautonomous gradient system (2.10); the existence of such a maximal solution is
guaranteed by Corollaries 2.9 and 2.8. Since $\nabla_{g} \mathscr{E}(u)$ and $f$ are continuous functions, any (maximal) solution of (2.10) is continuously differentiable and satisfies the differential equation (2.10) in the classical sense (and not only in the integrated sense from Corollary 2.9). We will show that $T_{\max }=T$.

We multiply equation (2.10) by $\dot{u}$ with respect to the inner product $\langle\cdot, \cdot\rangle_{g(u)}$, integrate the result over $[0, t]\left(0<t<T_{\max } \leq T\right)$, apply the Cauchy-Schwarz inequality and the assumption on the metric $g$ in order to obtain

$$
\begin{aligned}
\int_{0}^{t}\|\dot{u}\|_{g(u)}^{2}+\int_{0}^{t}\left\langle\nabla_{g} \mathscr{E}(u), \dot{u}\right\rangle_{g(u)} & =\int_{0}^{t}\langle f, \dot{u}\rangle_{g(u)} \\
& \leq \frac{1}{2} \int_{0}^{t}\|f\|_{g(u)}^{2}+\frac{1}{2} \int_{0}^{t}\|\dot{u}\|_{g(u)}^{2} \\
& \leq \frac{C}{2} \int_{0}^{T}\|f\|_{e u c}^{2}+\frac{1}{2} \int_{0}^{t}\|\dot{u}\|_{g(u)}^{2}
\end{aligned}
$$

By definition of the gradient and by the chain rule, we have $\left\langle\nabla_{g} \mathscr{E}(u), \dot{u}\right\rangle_{g(u)}=$ $\mathscr{E}^{\prime}(u) \dot{u}=\frac{d}{d t} \mathscr{E}(u)$. This and the preceding inequality imply that

$$
\frac{1}{2} \int_{0}^{t}\|\dot{u}\|_{g(u)}^{2}+\mathscr{E}(u(t)) \leq \mathscr{E}\left(u_{0}\right)+\frac{C}{2} \int_{0}^{T}\|f\|_{e u c}^{2}=: c<\infty
$$

The right-hand side of this inequality is independent of $t \in\left[0, T_{\max }\right.$ ), and the first term on the left-hand side is positive. As a consequence, the range $\{u(t): t \in$ $\left.\left[0, T_{\max }\right)\right\}$ is a subset of the sub-level set $K_{c}$. By coercivity and continuity of $\mathscr{E}$, $K_{c}$ is bounded and closed in $\mathbb{R}^{d}$ which is finite dimensional. Hence, $K_{c}$ is compact.

Since the gradient $\nabla_{g} \mathscr{E}$ is continuous on $K_{c}$ (Lemma 2.4) and since the range of $u$ is contained in $K_{c}$, we obtain $\sup _{t \in\left[0, T_{\text {max }}\right)}\left\|\nabla_{g} \mathscr{E}(u(t))\right\|_{e u c}<\infty$. The differential equation (2.10) then implies that $\|\dot{u}\|_{e u c}$ is bounded, and integrable, on $\left[0, T_{\max }\right)$. Hence, $u$ extends to a continuous function on the closed interval $\left[0, T_{\max }\right]$.

Now, if $T_{\max }<T$, then we could extend the solution $u$ to a larger interval by solving the gradient system $\dot{u}+\nabla_{g} \mathscr{E}(u)=f$ with initial time $T_{\max }$ and initial value $u\left(T_{\max }\right)$. This would contradict to the fact that $u$ is already a maximal solution, and hence $T_{\max }<T$ is not possible. As a consequence, $T_{\max }=T$, that is, $u$ is a global solution.

### 2.5 Newton's method

Let $U \subseteq \mathbb{R}^{d}$ be an open set, and let $F: U \rightarrow \mathbb{R}^{d}$ be a continuously differentiable function. Consider again the problem of finding a solution $\bar{u} \in U$ of the algebraic equation

$$
F(\bar{u})=0 .
$$

We repeat the idea from the previous lecture and consider the function $\mathscr{E}: U \rightarrow \mathbb{R}$ given by $\mathscr{E}(u)=\frac{1}{2}\|F(u)\|^{2}$, where $\|\cdot\|$ is the euclidean norm on $\mathbb{R}^{d}$. Then one has
$F(\bar{u})=0$ if and only if $\mathscr{E}(\bar{u})=0$, and one may try to find a solution of the latter problem by considering a gradient system associated with $\mathscr{E}$.

In this example, we assume that the derivative

$$
F^{\prime}(u) \text { is invertible for every } u \in U
$$

and we define a Riemannian metric $g: U \rightarrow \operatorname{Inner}\left(\mathbb{R}^{d}\right)$ by setting

$$
\langle v, w\rangle_{g(u)}=\left\langle F^{\prime}(u) v, F^{\prime}(u) w\right\rangle_{e u c}, \quad u \in U, v, w \in \mathbb{R}^{d} .
$$

For every $u \in U$ the bracket $\langle\cdot, \cdot\rangle_{g(u)}$ is clearly bilinear and symmetric. Moreover, for every $u \in U, v \in \mathbb{R}^{d}$,

$$
\begin{aligned}
\langle v, v\rangle_{g(u)} & =\left\langle F^{\prime}(u) v, F^{\prime}(u) v\right\rangle_{e u c} \\
& =\left\|F^{\prime}(u) v\right\|_{e u c}^{2} \geq 0,
\end{aligned}
$$

so that $\langle\cdot, \cdot\rangle_{g(u)}$ is in addition positive semidefinite. If $\langle v, v\rangle_{g(u)}=0$, then $\left\|F^{\prime}(u) v\right\|_{e u c}^{2}=0$ and hence $F^{\prime}(u) v=0$ by the definiteness of the euclidean inner product. Since $F^{\prime}(u)$ is invertible by assumption, this implies $v=0$, and therefore $\langle\cdot, \cdot\rangle_{g(u)}$ is definite, too. Continuity of $g$ follows from the continuity of $F^{\prime}$.

For the gradient of $\mathscr{E}$ with respect to the metric $g$ we have

$$
\begin{aligned}
\left\langle\nabla_{g} \mathscr{E}(u), v\right\rangle_{g(u)} & =\mathscr{E}^{\prime}(u) v & & \text { (definition of gradient) } \\
& =\left\langle F(u), F^{\prime}(u) v\right\rangle_{e u c} & & \text { (compute } \mathscr{E}^{\prime} \text { ) } \\
& =\left\langle F^{\prime}(u) F^{\prime}(u)^{-1} F(u), F^{\prime}(u) v\right\rangle_{e u c} & & \text { (insert the term } F^{\prime}(u) F^{\prime}(u)^{-1} \text { ) } \\
& =\left\langle F^{\prime}(u)^{-1} F(u), v\right\rangle_{g(u)} & & \text { (definition of metric } g \text { ). }
\end{aligned}
$$

Comparing left-hand side and right-hand side yields

$$
\nabla_{g} \mathscr{E}(u)=F^{\prime}(u)^{-1} F(u) .
$$

Hence, the differential equation

$$
\begin{equation*}
\dot{u}+F^{\prime}(u)^{-1} F(u)=0 \tag{2.11}
\end{equation*}
$$

is a gradient system associated with the energy $\mathscr{E}$. We call this differential equation continuous Newton's method since it is a continuous (in time) version of the discrete Newton algorithm

$$
\begin{equation*}
u_{n+1}-u_{n}+F^{\prime}\left(u_{n}\right)^{-1} F\left(u_{n}\right)=0, \quad n \geq 0 . \tag{2.12}
\end{equation*}
$$

Much could be said about this algorithm which Newton proposed in his De analysi per aequationes numero terminorum infinitas (1669) and in his Methods of series
and fluxions (1671) for a special polynomial $F^{1}$. It was originally not designed to be a steepest descent algorithm, that is, a discrete version of a gradient system. The original idea of Newton's method is very different and involves no energy function: given an approximation $u_{n}$ of the exact solution $\bar{u}$ of the equation $F(\bar{u})=0$, Taylor expansion of order 1 yields the equation

$$
0=F(\bar{u})=F\left(u_{n}\right)+F^{\prime}\left(u_{n}\right)\left(\bar{u}-u_{n}\right)+o\left(\bar{u}-u_{n}\right)
$$

Dropping the rest term $o\left(\bar{u}-u_{n}\right)$ by assuming that it is small compared to the other terms if $u_{n}$ is close to $\bar{u}$, and solving the resulting linear equation

$$
0=F\left(u_{n}\right)+F^{\prime}\left(u_{n}\right)\left(u_{n+1}-u_{n}\right)
$$

leads to the iteration formula (2.12). It turns out that if $u_{0}$ is an initial value close to $\bar{u}$ (close in a sense which can in general not be quantified), then the sequence $\left(u_{n}\right)$ of iterates given by the Newton algorithm converges to $\bar{u}$ as $n \rightarrow \infty$. We will not give a proof of this fact, but we rather content ourselves of having proved that continuous Newton's method is a gradient system ${ }^{2}$.

Proposition 2.11. Assume that $F^{\prime}(v)$ is invertible for every $v \in U$. Assume further that there exists a (global) solution $u: \mathbb{R}_{+} \rightarrow \mathbb{R}^{d}$ of continuous Newton's method (2.11), and assume that $u$ has relatively compact range in $U$. Then there exists $\bar{u} \in U$ such that $F(\bar{u})=0$.

Proof. See the proof of Proposition 1.7.

### 2.6 Exercises

2.1. Prove Corollary 2.9.
2.2. Show that a function $\mathscr{E}: \mathbb{R}^{d} \rightarrow \mathbb{R}$ is coercive if and only if $\lim _{\|u\| \rightarrow \infty} \mathscr{E}(u)=+\infty$.
2.3 (Damped pendulum). Given $\alpha>0, \beta>0, \varepsilon>0$, consider the two functions $G: \mathbb{R}^{2} \rightarrow \mathbb{R}^{2}$ and $\mathscr{E}: \mathbb{R}^{2} \rightarrow \mathbb{R}$ given by

$$
G\binom{u_{1}}{u_{2}}=\binom{-u_{2}}{\alpha u_{2}+\beta \sin u_{1}}
$$

and

[^4]$$
\mathscr{E}\binom{u_{1}}{u_{2}}=-\beta \cos u_{1}+\frac{1}{2} u_{2}^{2}+\varepsilon u_{2} \sin u_{1}
$$
a) Prove that for every $u \in \mathbb{R}^{2} \backslash\{(k \pi, 0): k \in \mathbb{Z}\}$ and every $\varepsilon>0$ small enough one has
$$
\left\langle\nabla_{\text {euc }} \mathscr{E}(u), G(u)\right\rangle_{\text {euc }}>0
$$

Hint: Recall the special case of Young's inequality, $a b \leq \frac{1}{2 \lambda} a^{2}+\frac{\lambda}{2} b^{2}$, which is true for every $a, b \geq 0, \lambda>0$.
b) For every $u \in \mathbb{R}^{2}$, consider the matrix

$$
A(u)=\left(\begin{array}{cc}
-u_{2}-\varepsilon \sin u_{1} & -u_{2} \\
\beta \sin u_{1}+\varepsilon u_{2} \cos u_{1} \alpha u_{2}+\beta \sin u_{1}
\end{array}\right)
$$

Prove that for every $u \in \mathbb{R}^{2} \backslash\{(k \pi, 0): k \in \mathbb{Z}\}$ and every $\varepsilon>0$ small enough the matrix $A(u)$ is invertible.
Hint: Calculate $\operatorname{det} A(u)$ and compare with (a).
c) Prove that for every $u \in \mathbb{R}^{2} \backslash\{(k \pi, 0): k \in \mathbb{Z}\}$ and every $\varepsilon>0$ small enough

$$
\langle v, w\rangle_{g(u)}:=\left\langle\nabla_{e u c} \mathscr{E}(u), G(u)\right\rangle_{e u c}\left\langle A(u)^{-1} v, A(u)^{-1} w\right\rangle_{e u c}, \quad v, w \in \mathbb{R}^{2}
$$

defines an inner product on $\mathbb{R}^{2}$. The function $g: \mathbb{R}^{2} \backslash\{(k \pi, 0): k \in \mathbb{Z}\} \rightarrow$ $\operatorname{Inner}\left(\mathbb{R}^{2}\right)$ thus defined is a Riemannian metric.
d) Show that (for $\varepsilon>0$ small enough) $\nabla_{g} \mathscr{E}=G$.
e) Conclude that the differential system modeling the damped pendulum

$$
\begin{aligned}
\dot{u}_{1}-u_{2} & =0 \\
\dot{u}_{2}+\alpha u_{2}+\beta \sin u_{1} & =0
\end{aligned}
$$

is a gradient system on $\mathbb{R}^{2} \backslash\{(k \pi, 0): k \in \mathbb{Z}\}$, that is, outside the set of equilibrium points.
2.4. Let $F: \mathbb{R}^{d} \rightarrow \mathbb{R}^{d}$ be a continuously differentiable function such that $F^{\prime}(u)$ is invertible for every $u \in \mathbb{R}^{d}$ and such that $\lim _{\|u\| \rightarrow \infty}\|F(u)\|=+\infty$.
a) Show that the equation $F(\bar{u})=0$ has a solution.
b) Show that $F$ is surjective.

Hint: Replace the function $F$ by $F-v$, where $v \in \mathbb{R}^{d}$ is an arbitrary given vector.

## Lecture 3

## Gradients in infinite dimensional space: the Laplace operator on a bounded interval

Starting with this lecture, we consider real-valued functions which are defined on infinite dimensional Banach spaces. We study their gradients and the associated gradient systems. Many of the aspects and objects which we know for gradient systems in finite dimensional space reappear in the following study: the notion of derivative, ambient inner product or metric, gradient, gradient system, energy function, dissipation. But the analysis of gradient systems in infinite dimensional spaces - local and global existence of solutions, for example - requires more effort.

The effort is worthwhile! Many elliptic partial differential operators which appear in evolution models from physics, biology, chemistry or engineering turn out to be gradients, and the corresponding evolution models like diffusion models, phase separation models are gradient systems.

The most prominent example of an elliptic operator is the Laplace operator

$$
\Delta=\operatorname{div} \nabla=\sum_{i=1}^{d} \frac{\partial^{2}}{\partial x_{i}^{2}}
$$

acting on functions defined on domains in $\mathbb{R}^{d}$. We show that the Laplace operator on a bounded interval (the second derivative!), equipped with Dirichlet boundary conditions, is a gradient. This fundamental observation is not only important for the study of the heat equation $u_{t}-u_{x x}=0$ or similar parabolic equations. It is also important for the study of elliptic boundary value problems by variational methods. These problems and the beautiful theory of the calculus of variations are touched very slightly in this and the following lecture. The study of gradient systems is postponed for a while.

In a first step, we define the gradient of a real-valued function defined on an infinite dimensional Banach space. Our first examples of gradients, in particular the Laplace operator and other elliptic partial differential operators, are gradients of
quadratic forms. We devote a short separate section to quadratic forms.
Throughout this lecture, $V$ is a real Banach space with norm $\|\cdot\|_{V}$. We denote by $V^{\prime}$ the dual space of $V$, that is, the space of all continuous linear functionals $V \rightarrow \mathbb{R}$. If $u^{\prime} \in V^{\prime}$, then we write

$$
u^{\prime}(u) \text { or } u^{\prime} u \text { or }\left\langle u^{\prime}, u\right\rangle \text { or }\left\langle u^{\prime}, u\right\rangle_{V^{\prime}, V}
$$

for the value of $u^{\prime}$ at the element $u \in V$. The dual space $V^{\prime}$ is equipped with the norm $\left\|u^{\prime}\right\|_{V^{\prime}}=\sup _{\|u\|_{V} \leq 1}\left\langle u^{\prime}, u\right\rangle$.

### 3.1 Definition of gradient

Let $U \subseteq V$ be an open subset of $V$, and let $\mathscr{E}: U \rightarrow \mathbb{R}$ be a function. We say that $\mathscr{E}$ is differentiable if for every $u \in U$ there exists a continuous linear functional $u^{\prime} \in V^{\prime}$ such that

$$
\lim _{\|h\|_{V} \rightarrow 0} \frac{\mathscr{E}(u+h)-\mathscr{E}(u)-\left\langle u^{\prime}, h\right\rangle}{\|h\|_{V}}=0 .
$$

The functional $u^{\prime} \in V^{\prime}$, if it exists, is unique. The derivative is the function $\mathscr{E}^{\prime}: U \rightarrow V^{\prime}$ which assigns to every $u \in U$ the unique linear functional $u^{\prime} \in V^{\prime}$ for which the above equality holds. Consequently, we denote the derivative of $\mathscr{E}$ at a point $u$ by $\mathscr{E}^{\prime}(u)$. We say that $\mathscr{E}$ is continuously differentiable if $\mathscr{E}$ is differentiable and if the derivative is continuous from $U$ into $V^{\prime}$. The set of all continuously differentiable functions $U \rightarrow \mathbb{R}$ is a vector space which is denoted by $C^{1}(U)$.

Let $H$ be a real Hilbert space with inner product $\langle\cdot, \cdot\rangle_{H}$ and associated norm $\|\cdot\|_{H}$. Assume that $V$ is continuously and densely embedded into $H$. This means that we can identify $V$ with a dense subspace of $H$ (via some injection $V \rightarrow H$ ) and there exists a constant $C \geq 0$ such that $\|u\|_{H} \leq C\|u\|_{V}$ for every $u \in V$. We will write $V \hookrightarrow H$ for this situation.

Given a differentiable function $\mathscr{E}: U \rightarrow \mathbb{R}$, we define the gradient $\nabla_{H} \mathscr{E}$ with respect to the inner product $\langle\cdot, \cdot\rangle_{H}$ by

$$
\begin{aligned}
& D\left(\nabla_{H} \mathscr{E}\right):=\{u \in U: \text { there exists } v \in H \text { such that for every } \\
& \nabla_{H} \mathscr{E}(u):=v,
\end{aligned}
$$

that is, $\nabla_{H} \mathscr{E}(u)$ is the unique element in $H$, if it exists, which represents the derivative $\mathscr{E}^{\prime}(u)$ with respect to the inner product in $H$ :

$$
\begin{equation*}
\mathscr{E}^{\prime}(u) \varphi=\left\langle\nabla_{H} \mathscr{E}(u), \varphi\right\rangle_{H} \quad \text { for every } \varphi \in V . \tag{3.1}
\end{equation*}
$$

The "if it exists" is important and marks a difference to gradients in finite dimensional space. The fact that we consider two (in general different) spaces $V$ and $H$ leads to the necessity to consider the domain $D\left(\nabla_{H} \mathscr{E}\right)$ of the gradient of $\nabla_{H} \mathscr{E}$. This domain is in general strictly contained in $U$; not every derivative $\mathscr{E}^{\prime}(u)$ can be represented by some element in $H$ (see the various examples in this lecture).

As soon as $V$ is densely and continuously embedded into $H$, the dual $H^{\prime}$ is continuously embedded into the dual $V^{\prime}$; the restriction to $V$ of a continuous linear functional $H \rightarrow \mathbb{R}$ defines a continuous linear functional $V \rightarrow \mathbb{R}$. The resulting operator $H^{\prime} \rightarrow V^{\prime}$ is linear and continuous, and it is injective by the fact that $V$ is dense in $H$. But the embedding $H^{\prime} \rightarrow V^{\prime}$ is in general not surjective, that is, not every bounded linear functional $u^{\prime} \in V^{\prime}$ extends to a bounded linear functional on $H$. As a consequence, it may happen that a derivative $\mathscr{E}^{\prime}(u) \in V^{\prime}$ does not extend to a linear functional on $H$.

However, if a bounded linear functional $u^{\prime} \in V^{\prime}$ (for example, the derivative $\mathscr{E}^{\prime}(u)$ ) extends to a bounded linear functional on $H$, that is, if $u^{\prime} \in H^{\prime}$, then this functional may be represented by some element $u \in H$ due to a generalization of the Representation Lemmas 1.1 and 2.1. The Riesz-Fréchet representation theorem (Theorem D.53) says that if $H$ is a Hilbert space, then for every bounded linear functional $u^{\prime} \in H^{\prime}$ there exists a unique element $u \in H$ such that

$$
\left\langle u^{\prime}, v\right\rangle_{H^{\prime}, H}=\langle u, v\rangle_{H} \text { for every } v \in H .
$$

By the Riesz-Fréchet representation theorem, an element $u \in U$ belongs to the domain $D\left(\nabla_{H} \mathscr{E}\right)$ if and only if the derivative $\mathscr{E}^{\prime}(u)$ extends to a continuous linear functional on $H$.

Similarly as above, one may define the gradient with respect to a metric. Given a Hilbert space $H$ with inner product $\langle\cdot, \cdot\rangle_{H}$, we let $\mathscr{L}_{2}(H ; \mathbb{R})$ be the space of all bounded bilinear forms $a: H \times H \rightarrow \mathbb{R}$, that is, the space of all forms such that for every $u, v, w \in H$ and every $\lambda \in \mathbb{R}$ one has

$$
\begin{aligned}
& a(\lambda u+v, w)=\lambda a(u, w)+a(v, w) \quad \text { and } \\
& a(u, \lambda v+w)=\lambda a(u, v)+a(u, w),
\end{aligned}
$$

and

$$
|a(u, v)| \leq C\|u\|_{H}\|v\|_{H} \quad \text { for every } u, v \in H \text { and some } C \geq 0
$$

This space is a vector space for the natural addition and the natural scalar multiplication, and it is a Banach space for the norm

$$
\|a\|:=\sup _{\substack{\|u\|_{H} \leq 1 \\\|v\|_{H} \leq 1}}|a(u, v)| .
$$

Besides the norm convergence we consider also strong convergence of sequences. We say that a sequence $\left(a_{n}\right) \subset \mathscr{L}_{2}(H ; \mathbb{R})$ converges strongly to an element $a \in$
$\mathscr{L}_{2}(H ; \mathbb{R})$ if for every $u, v \in H$ one has $\lim _{n \rightarrow \infty} a_{n}(u, v)=a(u, v)$. Every norm convergent sequence is strongly convergent.

The set of all inner products on $H$, which we denote by $\operatorname{Inner}(H)$, is a subset of $\mathscr{L}_{2}(H ; \mathbb{R})$. It is therefore natural to speak of norm convergent sequences and strongly convergent sequences in $\operatorname{Inner}(H)$. A metric on an open subset $U \subseteq V$ is a function $g: U \rightarrow \operatorname{Inner}(H)$ which maps norm convergent sequences into strongly convergent sequences, that is, whenever $\lim _{n \rightarrow \infty}\left\|u_{n}-u\right\|_{V}=0, u_{n}, u \in U$, then $\left(g\left(u_{n}\right)\right)$ converges strongly to $g(u)$. For every $u \in U$, the inner product $g(u)$ is also denoted by the bracket $\langle\cdot, \cdot\rangle_{g(u)}$, and the induced norm is denoted by $\|\cdot\|_{g(u)}$.

Given a differentiable function $\mathscr{E}: U \rightarrow \mathbb{R}$ and given a metric $g$, we define the gradient $\nabla_{g} \mathscr{E}$ with respect to $g$ by

$$
\begin{aligned}
& D\left(\nabla_{g} \mathscr{E}\right):=\{u \in U: \text { there exists } v \in H \text { such that for every } \\
& \left.\qquad \varphi \in V \text { one has } \mathscr{E}^{\prime}(u) \varphi=\langle v, \varphi\rangle_{g(u)}\right\}, \text { and } \\
& \nabla_{g} \mathscr{E}(u):=v,
\end{aligned}
$$

that is, $\nabla_{g} \mathscr{E}(u)$ is the unique element in $H$, if it exists, which represents the derivative $\mathscr{E}^{\prime}(u)$ with respect to the inner product $\langle\cdot, \cdot\rangle_{g(u)}$ :

$$
\mathscr{E}^{\prime}(u) \varphi=\left\langle\nabla_{g} \mathscr{E}(u), \varphi\right\rangle_{g(u)} \quad \text { for every } \varphi \in V
$$

If, for every $u \in U$, the inner product $\langle\cdot, \cdot\rangle_{g(u)}$ is equivalent to the inner product $\langle\cdot, \cdot\rangle_{H}$, then an element $u \in U$ belongs to the domain $D\left(\nabla_{g} \mathscr{E}\right)$ if and only if the derivative $\mathscr{E}^{\prime}(u)$ extends to a continuous linear functional on $H$.

### 3.2 Gradients of quadratic forms

A function $\mathscr{E}: V \rightarrow \mathbb{R}$ is a quadratic form if there exists a bilinear, symmetric form $a: V \times V \rightarrow \mathbb{R}$ such that

$$
\mathscr{E}(u)=\frac{1}{2} a(u, u) .
$$

Recall that a bilinear, symmetric form is a bilinear form $a: V \times V \rightarrow \mathbb{R}$ such that for every $u, v \in V$ one has $a(u, v)=a(v, u)$.

Proposition 3.1. A quadratic form $\mathscr{E}: V \rightarrow \mathbb{R}$ is continuous if and only if the associated bilinear form a is continuous.

Proof. Clearly, if a bilinear form $a$ is continuous, then the associated quadratic form $\mathscr{E}$ is continuous. In order to prove the converse, assume that the function $u \mapsto \frac{1}{2} a(u, u)$ is continuous. Then, by the polarization identity

$$
a(u, v)=\frac{1}{4}(a(u+v, u+v)-a(u-v, u-v))
$$

$a$ is continuous, too.
Proposition 3.2. If $\mathscr{E}: V \rightarrow \mathbb{R}$ is a continuous quadratic form with associated bilinear form $a: V \times V \rightarrow \mathbb{R}$, then $\mathscr{E}$ is continuously differentiable and

$$
\mathscr{E}^{\prime}(u) v=a(u, v) \quad \text { for every } u, v \in V .
$$

The proof of this proposition is left as an exercise, but we point out that actually more is true: every continuous quadratic form is infinitely many times continuously differentiable.

Given a continuous quadratic form $\mathscr{E}: V \rightarrow \mathbb{R}$ with associated bilinear form $a$ : $V \times V \rightarrow \mathbb{R}$, and given a Hilbert space $H$ with inner product $\langle\cdot, \cdot\rangle_{H}$, such that $V$ is densely and continuously embedded into $H$, we may compute the gradient of $\mathscr{E}$ with respect to the inner product in $H$ by using the definition of the gradient and Proposition 3.2. They imply that

$$
\begin{aligned}
& D\left(\nabla_{H} \mathscr{E}\right)=\{u \in V: \text { there exists } v \in H \text { such that for every } \\
& \left.\qquad \varphi \in V \text { one has } a(u, \varphi)=\langle v, \varphi\rangle_{H}\right\} \text { and } \\
& \nabla_{H} \mathscr{E}(u)=v .
\end{aligned}
$$

It turns out that the gradient of a quadratic form $\mathscr{E}$ with respect to an inner product $\langle\cdot, \cdot\rangle_{H}$ is a linear operator on $H$. The domain $D\left(\nabla_{H} \mathscr{E}\right)$ is a linear subspace of $H$ and the mapping $u \mapsto \nabla_{H} \mathscr{E}(u)$ is linear (exercise). We point out that the gradient of a quadratic form with respect to a metric is in general not linear.

Quadratic forms arise naturally in the study of linear elliptic operators. These elliptic operators, in turn, appear in many parabolic equations. As already indicated in the introduction of this lecture, we spend some time in order to introduce elliptic operators. This implies that we have to define Sobolev spaces on domains. To start with, we confine ourselves to the one-dimensional case.

### 3.3 Sobolev spaces in one space dimension

In this section we define Sobolev spaces on intervals. We will not prove any properties of Sobolev spaces, and actually, for this lecture it is not necessary to prove any properties. It suffices to know the definition of Sobolev spaces and the notion of weak derivative.

Let $I \subseteq \mathbb{R}$ be an open interval. For every continuous function $\varphi: I \rightarrow \mathbb{R}$ we define the support

$$
\operatorname{supp} \varphi:=\overline{\{x \in I: \varphi(x) \neq 0\}}
$$

where the closure is to be taken in $I$. The space

$$
\mathscr{D}(I):=C_{c}^{\infty}(I):=\left\{\varphi \in C^{\infty}(I): \operatorname{supp} \varphi \subset I \text { is compact }\right\}
$$

is called the space of all test functions on $I$.
By the fundamental rule of partial integration, if $u, v:[a, b] \rightarrow \mathbb{R}$ are continuously differentiable functions on some compact interval $[a, b]$, then

$$
\int_{a}^{b} u v^{\prime}=u(b) v(b)-u(a) v(a)-\int_{a}^{b} u^{\prime} v
$$

In particular, for every $u \in C^{1}([a, b])$ and every $\varphi \in \mathscr{D}(a, b)$

$$
\begin{equation*}
\int_{a}^{b} u \varphi^{\prime}=-\int_{a}^{b} u^{\prime} \varphi \tag{3.2}
\end{equation*}
$$

since $\varphi(a)=\varphi(b)=0$. In the following definition, this partial integration rule is not a theorem, but it serves for the definition of a derivative. For every $-\infty \leq a<b \leq \infty$ and every $1 \leq p \leq \infty$, we define the first Sobolev space

$$
\begin{aligned}
& W^{1, p}(a, b):=\left\{u \in L^{p}(a, b): \text { there exists } v \in L^{p}(a, b)\right. \text { such that for every } \\
& \left.\qquad \varphi \in \mathscr{D}(a, b) \text { one has } \int_{a}^{b} u \varphi^{\prime}=-\int_{a}^{b} v \varphi\right\}
\end{aligned}
$$

If $p=2$, then we also write $H^{1}(a, b):=W^{1,2}(a, b)$.
By Lemma G.4, the function $v \in L^{p}(a, b)$ is uniquely determined, if it exists. In the following, we write $u^{\prime}:=v$, in accordance with the partial integration rule (3.2), and we call $u^{\prime}$ the weak derivative of $u$. With this notation, the space $W^{1, p}(a, b)$ is the space of all functions in $L^{p}(a, b)$ which admit a weak derivative in $L^{p}(a, b)$. The spaces $W^{1, p}(a, b)$ are Banach spaces for the norms

$$
\begin{aligned}
\|u\|_{W^{1, p}} & :=\left(\|u\|_{L^{p}}^{p}+\left\|u^{\prime}\right\|_{L^{p}}^{p}\right)^{\frac{1}{p}}, \quad 1 \leq p<\infty, \text { and } \\
\|u\|_{W^{1, \infty}} & :=\sup \left\{\|u\|_{L^{\infty}},\left\|u^{\prime}\right\|_{L^{\infty}}\right\}
\end{aligned}
$$

and the space $H^{1}(a, b)$ is a Hilbert space for the inner product

$$
\langle u, v\rangle_{H^{1}}:=\int_{a}^{b} u v+\int_{a}^{b} u^{\prime} v^{\prime}
$$

We further define

$$
W_{0}^{1, p}(a, b):=\overline{\mathscr{D}(a, b)} \|^{\|\cdot\|_{W^{1, p}}},
$$

that is, $W_{0}^{1, p}(a, b)$ is the closure of the test functions in $W^{1, p}(a, b)$. The space $W_{0}^{1,2}(a, b)$ is also denoted $H_{0}^{1}(a, b)$. The following result about Sobolev spaces on the interval will be proved later.

Theorem 3.3. Let $(a, b)$ be a bounded interval, and let $1 \leq p \leq \infty$. Then:
3.4 The Dirichlet-Laplace operator on a bounded interval
a) Every function $u \in W^{1, p}(a, b)$ is continuous on $[a, b]$. More precisely: every function $u \in W^{1, p}(a, b)$ has a continuous representant $u:[a, b] \rightarrow \mathbb{R}$.
b) A function $u \in W^{1, p}(a, b)$ belongs to $W_{0}^{1, p}(a, b)$ if and only if $u(a)=u(b)=0$.

For every $1 \leq p \leq \infty$ and every $k \geq 2$ we define inductively the $k$-th Sobolev space

$$
W^{k, p}(a, b):=\left\{u \in W^{1, p}(a, b): u^{\prime} \in W^{k-1, p}(a, b)\right\}
$$

which is a Banach space for the norm

$$
\begin{aligned}
\|u\|_{W^{k, p}} & :=\left(\sum_{j=0}^{k}\left\|u^{(j)}\right\|_{L^{p}}^{p}\right)^{\frac{1}{p}} \quad \text { if } 1 \leq p<\infty, \text { or } \\
\|u\|_{W^{1, \infty}} & :=\sup \left\{\|u\|_{L^{\infty}},\left\|u^{\prime}\right\|_{L^{\infty}}, \ldots,\left\|u^{(k)}\right\|_{L^{\infty}}\right\} \quad \text { if } p=\infty .
\end{aligned}
$$

Here, $u^{(j)}$ denotes the $j$-th weak derivative of $u, u^{(0)}=u$. The space $H^{k}(a, b):=$ $W^{k, 2}(a, b)$ is a Hilbert space for the inner product

$$
\langle u, v\rangle_{H^{k}}:=\sum_{j=0}^{k}\left\langle u^{(j)}, v^{(j)}\right\rangle_{L^{2}} .
$$

### 3.4 The Dirichlet-Laplace operator on a bounded interval

We consider the Sobolev space $V=H_{0}^{1}(0,1)$, equipped with the $H^{1}$ norm which turns it into a Hilbert space. We consider in addition the quadratic form $\mathscr{E}$ : $H_{0}^{1}(0,1) \rightarrow \mathbb{R}$ given by

$$
\mathscr{E}(u)=\frac{1}{2} \int_{0}^{1}\left(u^{\prime}\right)^{2}, \quad u \in H_{0}^{1}(0,1)
$$

The associated bilinear form $a: H_{0}^{1}(0,1) \times H_{0}^{1}(0,1) \rightarrow \mathbb{R}$ is given by

$$
a(u, v)=\int_{0}^{1} u^{\prime} v^{\prime}, \quad u, v \in H_{0}^{1}(0,1)
$$

Both functions $\mathscr{E}$ and $a$ are continuous.
We consider the space $H=L^{2}(0,1)$ equipped with the usual inner product, that is

$$
\langle u, v\rangle_{L^{2}}=\int_{0}^{1} u v
$$

and compute the gradient of $\mathscr{E}$ with respect to this inner product. We only use the definition of the gradient, Proposition 3.2 and the definition of the Sobolev spaces $H^{1}(0,1)$ and $H^{2}(0,1)$. First, when we translate the definition of the gradient, we find

3 Gradients in infinite dimensional space: the Laplace operator on a bounded interval

$$
\begin{aligned}
& D\left(\nabla_{L^{2}} \mathscr{E}\right)=\left\{u \in H_{0}^{1}(0,1): \text { there exists } v \in L^{2}(0,1)\right. \text { such that for every } \\
& \left.\qquad \varphi \in H_{0}^{1}(0,1) \text { one has } \int_{0}^{1} u^{\prime} \varphi^{\prime}=\int_{0}^{1} v \varphi\right\} \\
& \nabla_{L^{2}} \mathscr{E}(u)=v
\end{aligned}
$$

This gradient may be characterized in a more explicit way, as follows.

Let $u \in D\left(\nabla_{L^{2} \mathscr{E}}\right)$. Then, by the above characterization of the domain, $u \in$ $H_{0}^{1}(0,1)$ and there exists $v \in L^{2}(0,1)$ such that for every $\varphi \in H_{0}^{1}(0,1)$ one has

$$
\int_{0}^{1} u^{\prime} \varphi^{\prime}=\int_{0}^{1} v \varphi=-\int_{0}^{1}(-v) \varphi
$$

In particular, this equality holds for every test function $\varphi \in \mathscr{D}(0,1)$. Recalling now the definition of the Sobolev space $H^{1}(0,1)$, we see that $u^{\prime}$ belongs to $H^{1}(0,1)$ (that is $u \in H^{2}(0,1)$ ) and $-u^{\prime \prime}=v$. This proves that $u \in H^{2}(0,1) \cap H_{0}^{1}(0,1)$ and $\nabla_{L^{2}} \mathscr{E}(u)=-u^{\prime \prime}$.

Conversely, let $u \in H^{2}(0,1) \cap H_{0}^{1}(0,1)$. Then $u^{\prime \prime} \in L^{2}(0,1), u^{\prime} \in H^{1}(0,1)$, and by definition of the Sobolev space $H_{0}^{1}(0,1)$ we obtain the equality $\int_{0}^{1} u^{\prime} \varphi^{\prime}=-\int_{0}^{1} u^{\prime \prime} \varphi$ for every test function $\varphi \in \mathscr{D}(0,1)$. By definition of the space $H_{0}^{1}(0,1)$, the test functions are dense in $H_{0}^{1}(0,1)$. It is then an exercise to show that the equality $\int_{0}^{1} u^{\prime} \varphi^{\prime}=-\int_{0}^{1} u^{\prime \prime} \varphi$ holds for every $\varphi \in H_{0}^{1}(0,1)$. Since $u^{\prime \prime} \in L^{2}(0,1)$, this proves $u \in D\left(\nabla_{L^{2}} \mathscr{E}\right)$.

Putting the above inclusions together, we have therefore proved that

$$
\begin{aligned}
D\left(\nabla_{L^{2}} \mathscr{E}\right) & =H^{2}(0,1) \cap H_{0}^{1}(0,1) \quad \text { and } \\
\nabla_{L^{2}} \mathscr{E}(u) & =-u^{\prime \prime}
\end{aligned}
$$

We set ${ }_{(0,1)}^{D} \Delta:=-\nabla_{L^{2}} \mathscr{E}$ and we call this negative gradient the Dirichlet-Laplace operator on the interval $(0,1)$. Solving the boundary value problem

$$
\left\{\begin{array}{l}
u \in H^{2}(0,1)  \tag{3.3}\\
-u^{\prime \prime}(x)=f(x) \quad \text { for } x \in(0,1) \\
u(0)=u(1)=0
\end{array}\right.
$$

for a given $f \in L^{2}(0,1)$ is equivalent to solving the abstract equation

$$
\nabla_{L^{2}} \mathscr{E}(u)=f \quad\left(\text { or: }-\frac{D}{(0,1)} \Delta u=f\right)
$$

in particular because the domain of the gradient coincides exactly with the space of all functions in $u \in H^{2}(0,1)$ which satisfy the Dirichlet boundary condition $u(0)=$ $u(1)=0$.

### 3.5 The Dirichlet-Laplace operator with multiplicative coefficient

We consider, as in the previous example, the Sobolev space $V=H_{0}^{1}(0,1)$ and the function $\mathscr{E}: H_{0}^{1}(0,1) \rightarrow \mathbb{R}$ given by

$$
\mathscr{E}(u)=\frac{1}{2} \int_{0}^{1}\left(u^{\prime}\right)^{2}, \quad u \in H_{0}^{1}(0,1)
$$

We also consider, as in the previous example, the space $H=L^{2}(0,1)$, but now it is equipped with the inner product

$$
\langle v, w\rangle=\int_{0}^{1} v w \frac{1}{m}
$$

where $m \in L^{\infty}(0,1)$ is a given positive function such that $\frac{1}{m} \in L^{\infty}(0,1)$. This inner product is equivalent to the usual inner product on $L^{2}(0,1)$, as one easily verifies. What is the gradient of $\mathscr{E}$ with respect to this new inner product?

As in the previous example, we first translate the definition of the gradient into this situation, by using Proposition 3.2, and we now find

$$
\begin{aligned}
& D(\nabla \mathscr{E})=\left\{u \in H_{0}^{1}(0,1): \text { there exists } v \in L^{2}(0,1)\right. \text { such that for every } \\
& \left.\qquad \varphi \in H_{0}^{1}(0,1) \text { one has } \int_{0}^{1} u^{\prime} \varphi^{\prime}=\int_{0}^{1} v \varphi \frac{1}{m}\right\}, \\
& \nabla \mathscr{E}(u)=v .
\end{aligned}
$$

This gradient may be characterized in a more explicit way, as follows.
Let $u \in D(\nabla \mathscr{E})$. Then $u \in H_{0}^{1}(0,1)$ and there exists $v \in L^{2}(0,1)$ such that for every $\varphi \in H_{0}^{1}(0,1)$

$$
\int_{0}^{1} u^{\prime} \varphi^{\prime}=\int_{0}^{1} v \varphi \frac{1}{m}=-\int_{0}^{1}\left(-v \frac{1}{m}\right) \varphi
$$

In particular, this equality holds for every test function $\varphi \in \mathscr{D}(0,1)$. Recalling the definition of the Sobolev space $H^{1}(0,1)$, we see that $u^{\prime}$ belongs to $H^{1}(0,1)$ (that is, $\left.u \in H^{2}(0,1)\right)$ and $-u^{\prime \prime}=v \frac{1}{m}$. This proves $u \in H^{2}(0,1) \cap H_{0}^{1}(0,1)$ and $\nabla \mathscr{E}(u)=$ $-m u^{\prime \prime}$.

Conversely, let $u \in H^{2}(0,1) \cap H_{0}^{1}(0,1)$. Then $u^{\prime \prime} \in L^{2}(0,1), u^{\prime} \in H^{1}(0,1)$, and by definition of the Sobolev space $H^{1}(0,1)$, we obtain the equality $\int_{0}^{1} u^{\prime} \varphi^{\prime}=$ $-\int_{0}^{1} u^{\prime \prime} \varphi=-\int_{0}^{1} m u^{\prime \prime} \varphi \frac{1}{m}$ first for every $\varphi \in \mathscr{D}(0,1)$, and then, since $\mathscr{D}(0,1)$ is dense in $H_{0}^{1}(0,1)$, also for every $\varphi \in H_{0}^{1}(0,1)$. Since $m u^{\prime \prime} \in L^{2}(0,1)$, this proves $u \in D(\nabla \mathscr{E})$.

Putting the above inclusions together, we have therefore proved that

$$
\begin{aligned}
D(\nabla \mathscr{E}) & =H^{2}(0,1) \cap H_{0}^{1}(0,1) \quad \text { and } \\
\nabla \mathscr{E}(u) & =-m u^{\prime \prime}
\end{aligned}
$$

Hence, solving the abstract equation

$$
\nabla \mathscr{E}(u)=f
$$

is now equivalent to solving the boundary value problem

$$
\left\{\begin{array}{l}
u \in H^{2}(0,1) \\
-m(x) u^{\prime \prime}(x)=f(x) \text { for } x \in(0,1) \\
u(0)=u(1)=0
\end{array}\right.
$$

We consider also the following nonlinear variant of the above example. Let $\varepsilon \in$ $(0,1)$, and let $m: \mathbb{R} \rightarrow\left[\varepsilon, \frac{1}{\varepsilon}\right]$ be a continuous function. For every $u \in H_{0}^{1}(0,1)$ we consider the inner product

$$
\langle v, w\rangle_{g(u)}=\int_{0}^{1} v w \frac{1}{m(u)}
$$

where $m(u)=m \circ u$ is the composition of the functions $m$ and $u$. The resulting function $g: H_{0}^{1}(0,1) \rightarrow \operatorname{Inner}\left(L^{2}(0,1)\right), u \mapsto\langle\cdot, \cdot\rangle_{g(u)}$ maps norm convergent sequences into strongly convergent sequences (exercise) and is therefore a metric. By repeating the above computation, one sees that the gradient of $\mathscr{E}$ with respect to the metric $g$ is given by

$$
\begin{aligned}
D\left(\nabla_{g} \mathscr{E}\right) & =H^{2}(0,1) \cap H_{0}^{1}(0,1) \quad \text { and } \\
\nabla_{g} \mathscr{E}(u) & =-m(u) u^{\prime \prime}
\end{aligned}
$$

Hence, solving the abstract equation

$$
\nabla_{g} \mathscr{E}(u)=f
$$

is now equivalent to solving the boundary value problem

$$
\left\{\begin{array}{l}
u \in H^{2}(0,1)  \tag{3.4}\\
-m(u(x)) u^{\prime \prime}(x)=f(x) \text { for } x \in(0,1) \\
u(0)=u(1)=0
\end{array}\right.
$$

This problem is in general nonlinear.

### 3.6 Exercises

### 3.1. Prove Proposition 3.2.

3.2. Show that the gradient of a quadratic function $\mathscr{E}: V \rightarrow \mathbb{R}$ with respect to an inner product $\langle\cdot, \cdot\rangle_{H}$ is a linear operator, that is, the domain $D\left(\nabla_{H} \mathscr{E}\right)$ is a linear subspace of $H$ and the mapping $u \mapsto \nabla_{H} \mathscr{E}(u)$ is linear.
3.3 (The Dirichlet-Laplace operator with first order coefficient). Let $b:[0,1] \rightarrow$ $\mathbb{R}$ be a continuous function and let $B(x):=\int_{0}^{x} b(y) d y$. Consider the energy function

$$
\begin{aligned}
\mathscr{E}: H_{0}^{1}(0,1) & \rightarrow \mathbb{R}, \\
u & \mapsto \frac{1}{2} \int_{0}^{1}\left(u^{\prime}\right)^{2} e^{-B} .
\end{aligned}
$$

Compute the gradient $\nabla_{H} \mathscr{E}$ with respect to the inner product $\langle v, w\rangle_{H}=\int_{0}^{1} v w e^{-B}$ on $H=L^{2}(0,1)$. More precisely, show that

$$
\begin{aligned}
D\left(\nabla_{H} \mathscr{E}\right) & =H^{2}(0,1) \cap H_{0}^{1}(0,1) \\
\nabla_{H} \mathscr{E}(u) & =-u^{\prime \prime}+b u^{\prime}
\end{aligned}
$$

3.4 (A second order operator in "divergence form"). Let $a \in L^{\infty}(0,1)$. Consider the energy function

$$
\begin{aligned}
\mathscr{E}: H_{0}^{1}(0,1) & \rightarrow \mathbb{R}, \\
u & \mapsto \frac{1}{2} \int_{0}^{1} a \cdot\left(u^{\prime}\right)^{2} .
\end{aligned}
$$

Compute the gradient $\nabla_{L^{2}} \mathscr{E}$ with respect to the usual inner product on $L^{2}(0,1)$, $\langle v, w\rangle_{L^{2}}=\int_{0}^{1} v w$. More precisely, show that

$$
\begin{aligned}
D\left(\nabla_{L^{2}} \mathscr{E}\right) & =\left\{u \in H_{0}^{1}(0,1): a u^{\prime} \in H^{1}(0,1)\right\} \\
\nabla_{L^{2}} \mathscr{E}(u) & =-\left(a u^{\prime}\right)^{\prime}
\end{aligned}
$$

3.5. Let $\varepsilon \in(0,1)$, and let $m: \mathbb{R} \rightarrow\left[\varepsilon, \frac{1}{\varepsilon}\right]$ be a continuous function. For every $u \in$ $H_{0}^{1}(0,1)$ we consider the inner product

$$
\langle v, w\rangle_{g(u)}=\int_{0}^{1} v w \frac{1}{m(u)}
$$

Show that the function $g: H_{0}^{1}(0,1) \rightarrow \operatorname{Inner}\left(L^{2}(0,1)\right), u \mapsto\langle\cdot, \cdot\rangle_{g(u)}$ is a metric, that is, it maps norm convergent sequences into strongly convergent sequences.

Hint. One may use, without proof, the Sobolev embedding theorem (Theorem G.14, Theorem 3.3) which says that the space $H_{0}^{1}(0,1)$ is continuously embedded into $C([0,1])$, that is, every function $u \in H_{0}^{1}(0,1)$ is continuous on $[0,1]$ and there exists a constant $C>0$ (independent of $u$ ) such that $\|u\|_{L^{\infty}} \leq C\|u\|_{H_{0}^{1}}$.

## Lecture 4

## Gradients in infinite dimensional space: the Laplace operator and the $p$-Laplace operator

In this lecture, we present further examples of gradients in infinite dimensional space: given an open set $\Omega \subseteq \mathbb{R}^{d}$, we define the Dirichlet-Laplace operator on $L^{2}(\Omega)$ and the Dirichlet- $p$-Laplace operator on $L^{2}(\Omega)$. The reason, why we give these operators so much importance, can only be given later, when we will discuss linear and nonlinear diffusion equations as examples of gradient systems. But actually the Laplace operator and the $p$-Laplace operator appear not only in diffusion equations. They also appear in many other models from mathematical physics, such as wave propagation models, models from fluid mechanics, or in elasticity.

In order to introduce the new examples, we need to define Sobolev spaces on open sets $\Omega \subseteq \mathbb{R}^{d}$. In this lecture, we write also

$$
u v:=\langle u, v\rangle_{e u c}=\sum_{i=1}^{d} u_{i} v_{i} \quad \text { for the euclidean inner product, and }
$$

$|u|$ for the euclidean norm on $\mathbb{R}^{d}$.
For a differentiable function $u: \Omega \rightarrow \mathbb{R}$, where $\Omega \subseteq \mathbb{R}^{d}$ is an open set, we denote by $\nabla u$ its euclidean gradient.

### 4.1 Sobolev spaces in higher dimensions

Let $\Omega \subseteq \mathbb{R}^{d}$ be an open set with boundary $\partial \Omega$. For every continuous function $\varphi$ : $\Omega \rightarrow \mathbb{R}$ we define the support

$$
\operatorname{supp} \varphi:=\overline{\{x \in \Omega: \varphi(x) \neq 0\}}
$$

where the closure is to be taken in $\Omega$. Let

$$
C_{c}^{1}(\Omega):=\left\{\varphi \in C^{1}(\Omega): \operatorname{supp} \varphi \text { is compact }\right\}
$$

The following theorem, the integration by parts in higher dimensions, is a consequence of the Divergence Theorem (Theorem F.6). In this theorem, $C^{1}(\bar{\Omega})$ is the set of all continuously differentiable functions $u: \Omega \rightarrow \mathbb{R}$ such that $u$ and $\nabla u$ admit continuous extensions to the closure $\bar{\Omega}$. For the definition of what we mean by an open set of class $C^{1}$ (a notion related to the regularity of the boundary), and the definition of the outer normal vector, we refer to Appendix F. Domains with "smooth" boundary (balls, for example) are of class $C^{1}$, while domains with corners (triangles, rectangles) or slits are not of class $C^{1}$.


A $C^{1}$ domain with some outer normal vectors


A domain with corners is not of class $C^{1}$


A domain with a slit (here an ellipse minus a curve) is not of class $C^{1}$.

Theorem 4.1 (Integration by parts). Let $\Omega \subseteq \mathbb{R}^{d}$ be open, bounded and of class $C^{1}$. Then there exists a unique Borel measure $\sigma$ on $\partial \Omega$ (called the surface measure) such that for every $u, v \in C^{1}(\bar{\Omega})$ and every $1 \leq i \leq d$

$$
\int_{\Omega} u \frac{\partial v}{\partial x_{i}}=\int_{\partial \Omega} u v n_{i} d \sigma-\int_{\Omega} \frac{\partial u}{\partial x_{i}} v
$$

where $n(x)=\left(n_{i}(x)\right)_{1 \leq i \leq d}$ denotes the outer normal vector at a point $x \in \partial \Omega$.
In particular, if $u \in C^{1}(\bar{\Omega})$ and $\varphi \in C_{c}^{1}(\Omega)$, then

$$
\int_{\Omega} u \frac{\partial \varphi}{\partial x_{i}}=-\int_{\Omega} \frac{\partial u}{\partial x_{i}} \varphi
$$

and one may prove that this equality holds without any regularity assumption on $\Omega$. Similarly as in the case of Sobolev spaces on intervals and weak derivatives of functions of one variable, the rule of integration by parts serves for the definition of weak partial derivatives. For every $1 \leq p \leq \infty$ and every open $\Omega \subseteq \mathbb{R}^{d}$ we define the first Sobolev space

$$
\begin{aligned}
W^{1, p}(\Omega):=\left\{u \in L^{p}(\Omega):\right. & \text { for every } i=1, \ldots, d \text { there exists } g_{i} \in L^{p}(\Omega) \text { such } \\
& \text { that for every } \left.\varphi \in C_{c}^{1}(\Omega): \int_{\Omega} u \frac{\partial \varphi}{\partial x_{i}}=-\int_{\Omega} g_{i} \varphi\right\}
\end{aligned}
$$

We also write $H^{1}(\Omega)$ instead of $W^{1,2}(\Omega)$. By Lemma G.4, the functions $g_{i} \in L^{p}(\Omega)$ are uniquely determined, if they exist. In the following, we write $\frac{\partial u}{\partial x_{i}}:=g_{i}$, and we
call $\frac{\partial u}{\partial x_{i}}$ the $i$-th weak partial derivative of $u$ and $\nabla u:=\left(\frac{\partial u}{\partial x_{1}}, \ldots, \frac{\partial u}{\partial x_{d}}\right)$ the weak euclidean gradient of $u$. With these definitions, the Sobolev space $W^{1, p}(\Omega)$ is the space of all functions in $L^{p}(\Omega)$ which admit all weak partial derivatives in $L^{p}(\Omega)$. If $u$ is continuously differentiable, then partial derivatives coincide with weak partial derivatives by Gauß' theorem, and the euclidean gradient coincides with the weak euclidean gradient by Lemma 1.2.

The Sobolev spaces $W^{1, p}(\Omega)$ are Banach spaces for the norms

$$
\begin{aligned}
\|u\|_{W^{1, p}} & :=\left(\|u\|_{L^{p}}^{p}+\sum_{i=1}^{d}\left\|\frac{\partial u}{\partial x_{i}}\right\|_{L^{p}}^{p}\right)^{\frac{1}{p}}, \quad 1 \leq p<\infty, \text { and } \\
\|u\|_{W^{1, \infty}} & :=\sup \left\{\|u\|_{L^{\infty}},\left\|\frac{\partial u}{\partial x_{1}}\right\|_{L^{\infty}}, \ldots,\left\|\frac{\partial u}{\partial x_{d}}\right\|_{L^{\infty}}\right\},
\end{aligned}
$$

and $H^{1}(\Omega)$ is a Hilbert space for the inner product

$$
\langle u, v\rangle_{H^{1}}:=\langle u, v\rangle_{L^{2}}+\sum_{i=1}^{d}\left\langle\frac{\partial u}{\partial x_{i}}, \frac{\partial v}{\partial x_{i}}\right\rangle_{L^{2}} .
$$

We define

$$
W_{0}^{1, p}(\Omega):=\overline{C_{c}^{1}(\Omega)} \|^{\|\cdot\|_{W^{1, p}}}
$$

that is, $W_{0}^{1, p}(\Omega)$ is the closure of the space $C_{c}^{1}(\Omega)$ in $W^{1, p}(\Omega)$, and we put $H_{0}^{1}(\Omega):=W_{0}^{1,2}(\Omega)$.

Not all properties of Sobolev spaces on intervals carry over to Sobolev spaces on open sets $\Omega \subseteq \mathbb{R}^{d}$. For example, it is not true that every function $u \in W^{1, p}(\Omega)$ is continuous on the closure $\bar{\Omega}$, or even in the interior $\Omega$, unless one makes further assumptions on $p$ and $\Omega$ (dimension of $\Omega$, regularity of the boundary)! It is therefore not immediately clear how to give a meaning to the restriction to the boundary $\partial \Omega$ of a function in $W^{1, p}(\Omega)$, since the boundary $\partial \Omega$ is in general a set of Lebesgue measure zero. The problem of defining such restrictions to the boundary - one calls them also traces - is not discussed here. We only say that it is possible to define such traces on the boundary, provided the boundary is regular enough. If the boundary $\partial \Omega$ is regular enough, then one can show that the space $W_{0}^{1, p}(\Omega)$ is the space of all functions in $W^{1, p}(\Omega)$ whose traces on the boundary $\partial \Omega$ vanish.

For every $1 \leq p \leq \infty$ and every $k \geq 2$ we define inductively the $k$-th Sobolev space

$$
W^{k, p}(\Omega):=\left\{u \in W^{1, p}(\Omega): \text { for every } i=1, \ldots, d \text { one has } \frac{\partial u}{\partial x_{i}} \in W^{k-1, p}(\Omega)\right\}
$$

which is a Banach space for the norm

$$
\begin{aligned}
\|u\|_{W^{k, p}} & :=\left(\|u\|_{L^{p}}^{p}+\sum_{i=1}^{d}\left\|\frac{\partial u}{\partial x_{i}}\right\|_{W^{k-1, p}}^{p}\right)^{\frac{1}{p}} \quad \text { if } 1 \leq p<\infty, \text { and } \\
\|u\|_{W^{k, \infty}} & :=\sup \left\{\|u\|_{L^{\infty}},\left\|\frac{\partial u}{\partial x_{1}}\right\|_{W^{k-1, \infty}}, \ldots,\left\|\frac{\partial u}{\partial x_{d}}\right\|_{W^{k-1, \infty}}\right\} \quad \text { if } p=\infty
\end{aligned}
$$

The space $H^{k}(\Omega):=W^{k, 2}(\Omega)$ is a Hilbert space for the inner product

$$
\langle u, v\rangle_{H^{k}}:=\langle u, v\rangle_{L^{2}}+\sum_{i=1}^{d}\left\langle\frac{\partial u}{\partial x_{i}}, \frac{\partial v}{\partial x_{i}}\right\rangle_{H^{k-1}}
$$

For further results about Sobolev spaces, we refer to the Appendix or to [Adams (1975)], [Adams and Fournier (2003)].

### 4.2 The Dirichlet-Laplace operator

Let $\Omega \subseteq \mathbb{R}^{d}$ be an open set and consider the Banach space $V=H_{0}^{1}(\Omega)$. We consider the function $\mathscr{E}: H_{0}^{1}(\Omega) \rightarrow \mathbb{R}$ given by

$$
\mathscr{E}(u)=\frac{1}{2} \int_{\Omega}|\nabla u|^{2}, \quad u \in H_{0}^{1}(\Omega) .
$$

This function is a quadratic form with associated bilinear form $a$ given by

$$
a(u, v)=\int_{\Omega} \nabla u \nabla v, \quad u, v \in H_{0}^{1}(\Omega)
$$

We consider the Hilbert space $H=L^{2}(\Omega)$, equipped with the usual inner product. Then $V$ is densely and continuously embedded into $H$ (for the density, it suffices to prove that $C_{c}^{1}(\Omega)$ is dense in $L^{2}(\Omega)$, see Theorem G.3). The gradient of $\mathscr{E}$ with respect to the usual $L^{2}$ inner product is, as computed in Section 3.2, given by

$$
\begin{aligned}
& D\left(\nabla_{L^{2}} \mathscr{E}\right)=\left\{u \in H_{0}^{1}(\Omega): \text { there exists } v \in L^{2}(\Omega)\right. \text { such that for every } \\
& \left.\qquad \varphi \in H_{0}^{1}(\Omega) \text { one has } \int_{\Omega} \nabla u \nabla \varphi=\int_{\Omega} v \varphi\right\} \text { and } \\
& \nabla_{L^{2}} \mathscr{E}(u)=v .
\end{aligned}
$$

We write ${ }_{\Omega}^{D} \Delta:=-\nabla_{L^{2}} \mathscr{E}$ and we call this negative gradient the Dirichlet-Laplace operator on $L^{2}(\Omega)$. This term is justified by the following facts.

First, the Laplace operator is the partial differential operator

$$
\Delta=\operatorname{div} \nabla=\sum_{i=1}^{d} \frac{\partial^{2}}{\partial x_{i}^{2}}
$$

Second, if $u \in H^{2}(\Omega)$, then, by definition, all weak partial derivatives $\frac{\partial u}{\partial x_{i}}, \frac{\partial^{2} u}{\partial x_{i} \partial x_{j}}$ exist and belong to $L^{2}(\Omega)$. In particular, $\Delta u \in L^{2}(\Omega)$, if we understand by $\Delta u$ the sum of the weak partial derivatives $\frac{\partial^{2} u}{\partial x_{i}^{2}}$. By definition of the weak derivatives (or: by definition of the Sobolev spaces $H^{1}$ and $H^{2}$ ), for every function $\varphi \in C_{c}^{1}(\Omega)$

$$
\begin{aligned}
\int_{\Omega} \nabla u \nabla \varphi & =\int_{\Omega} \sum_{i=1}^{d} \frac{\partial u}{\partial x_{i}} \frac{\partial \varphi}{\partial x_{i}}=\sum_{i=1}^{d} \int_{\Omega} \frac{\partial u}{\partial x_{i}} \frac{\partial \varphi}{\partial x_{i}} \\
& =-\sum_{i=1}^{d} \int_{\Omega} \frac{\partial^{2} u}{\partial x_{i}^{2}} \varphi=-\int_{\Omega} \sum_{i=1}^{d} \frac{\partial^{2} u}{\partial x_{i}^{2}} \varphi \\
& =-\int_{\Omega} \Delta u \varphi
\end{aligned}
$$

Since, by definition, the space $C_{c}^{1}(\Omega)$ is dense in $H_{0}^{1}(\Omega)$, the equality $\int_{\Omega} \nabla u \nabla \varphi=$ $-\int_{\Omega} \Delta u \varphi$ actually holds for every $\varphi \in H_{0}^{1}(\Omega)$. As a consequence of this and the definition of ${ }_{\Omega}^{D} \Delta$, we obtain that every $u \in H^{2}(\Omega) \cap H_{0}^{1}(\Omega)$ belongs to the domain $D\left({ }_{\Omega}^{D} \Delta\right)$ and

$$
{ }_{\Omega}^{D} \Delta u=\Delta u .
$$

Third, since $D\left({ }_{\Omega}^{D} \Delta\right) \subseteq H_{0}^{1}(\Omega)$, every function in the domain of the DirichletLaplace operator satisfies the Dirichlet boundary condition $\left.u\right|_{\partial \Omega}=0$ in the sense that its trace (the restriction of the function to the boundary, in the weak sense of Sobolev spaces) vanishes.

The Dirichlet-Laplace operator is of importance when one studies the boundary value problem

$$
\left\{\begin{aligned}
-\Delta u & =f \text { in } \Omega \\
u & =0 \text { on } \partial \Omega
\end{aligned}\right.
$$

for some given function $f \in L^{2}(\Omega)$. We call a function $u \in H_{0}^{1}(\Omega)$ a weak solution of this boundary value problem if for every $\varphi \in H_{0}^{1}(\Omega)$

$$
\int_{\Omega} \nabla u \nabla \varphi=\int_{\Omega} f \varphi
$$

Moreover, we call a function $u \in C^{2}(\bar{\Omega})$ a classical solution if it satisfies the equation $-\Delta u=f$ in $\Omega$ (all second derivatives being now classical derivatives) and the boundary condition $u=0$ on $\partial \Omega$, pointwise everywhere. By Gauß' theorem, every classical solution is a weak solution. Moreover, one shows in an exercise that $u$ is a weak solution if and only if

$$
-{ }_{\Omega}^{D} \Delta u=f
$$

Remark. Unfortunately and unlikely to the case of the Dirchlet-Laplace operator on the interval, we can not characterize the domain of the Dirichlet-Laplace operator on $\Omega$, if $\Omega \subseteq \mathbb{R}^{d}$ is a general open set. For example, one can not prove that $D\left({ }_{\Omega}^{D} \Delta\right)=$ $H^{2}(\Omega) \cap H_{0}^{1}(\Omega)$; only the inclusion $\supseteq$ is true in general, as we have shown above.

As soon as the dimension $d \geq 2$, one can not conclude from ${ }_{\Omega}^{D} \Delta u \in L^{2}(\Omega)$ that each weak partial derivative $\frac{\partial^{2} u}{\partial x_{i}^{2}}$ exists and belongs to $L^{2}(\Omega)$, not to speak of the weak mixed partial derivatives $\frac{\partial^{2} u}{\partial x_{i} \partial x_{j}}$. In fact, this is a difficult problem from the theory of elliptic operators, and we refer to [Gilbarg and Trudinger (2001), Chapter 8], for example. Only two remarks, without proof: first, if $\Omega^{\prime} \subset \Omega$ is an open subset such that the closure $\bar{\Omega}^{\prime}$ is still contained in $\Omega$, then the restriction of $u \in D\left({ }_{\Omega}^{D} \Delta\right)$ to $\Omega^{\prime}$ belongs to $H^{2}\left(\Omega^{\prime}\right)$ ("interior regularity"). Second, if $\Omega$ is of class $C^{2}$ (a regularity property of the boundary again), then one does have $D\left({ }_{\Omega}^{D} \Delta\right)=H^{2}(\Omega) \cap H_{0}^{1}(\Omega)$ ("global regularity"); for example, this equality is true if $\Omega$ is a ball. These results are not trivial and we summarize that the equality $D\left({ }_{\Omega}^{D} \Delta\right)=H^{2}(\Omega) \cap H_{0}^{1}(\Omega)$ is false in general.

### 4.3 The Dirichlet- $p$-Laplace operator

Let $\Omega \subseteq \mathbb{R}^{d}$ be an open set and let $p>1$. We consider the Banach space $V=$ $W_{0}^{1, p}(\Omega)$ and the function $\mathscr{E}: W_{0}^{1, p}(\Omega) \rightarrow \mathbb{R}$ given by

$$
\mathscr{E}(u)=\frac{1}{p} \int_{\Omega}|\nabla u|^{p}, \quad u \in W_{0}^{1, p}(\Omega) .
$$

This function is not a quadratic form (unless $p=2$ ), but still we can show that it is continuously differentiable and compute its derivative, by applying a general result which we state and prove only in the following section.

Theorem 4.2. The function $\mathscr{E}$ is continuously differentiable, and

$$
\begin{equation*}
\mathscr{E}^{\prime}(u) v=\int_{\Omega}|\nabla u|^{p-2} \nabla u \nabla v \quad \text { for every } u, v \in W_{0}^{1, p}(\Omega) \tag{4.1}
\end{equation*}
$$

Moreover, $\mathscr{E}$ and $\mathscr{E} \prime$ map bounded sets into bounded sets.
Proof. Consider the function $F: \Omega \times \mathbb{R}^{d} \rightarrow \mathbb{R}$ given by $F(x, u)=\frac{1}{p}|u|^{p}$. Then $F(x, 0)=0$, and the function $F$ is continuously differentiable. If $\frac{\partial F}{\partial u}(x, u)$ denotes the derivative of the function $F(x, \cdot)$ at $u \in \mathbb{R}^{d}$, then $\frac{\partial F}{\partial u}(x, u) v=|u|^{p-2} u v$ and $\left|\frac{\partial F}{\partial u}(x, u)\right|=|u|^{p-1}$ for every $u, v \in \mathbb{R}^{d}$. Therefore, the function $F$ satisfies the conditions of Theorem 4.3 below. By Theorem 4.3, the function $\mathscr{F}: L^{p}\left(\Omega ; \mathbb{R}^{d}\right) \rightarrow \mathbb{R}$ given by $\mathscr{F}(u)=\int_{\Omega} F(x, u(x)) d x$ is continuously differentiable and

$$
\mathscr{F}^{\prime}(u) v=\int_{\Omega}|u|^{p-2} u v \quad \text { for every } u, v \in L^{p}\left(\Omega ; \mathbb{R}^{d}\right)
$$

Now the statement of the theorem follows by observing that the function $\mathscr{E}$ is the composition of the function $\mathscr{F}$ above and the continuous, linear mapping $\mathscr{L}: W_{0}^{1, p}(\Omega) \rightarrow L^{p}\left(\Omega ; \mathbb{R}^{d}\right), u \mapsto \nabla u$, that is, $\mathscr{E}=\mathscr{F} \circ \mathscr{L}$. Every continuous lin-
ear mapping is continuously differentiable, so that $\mathscr{E}$ is continuously differentiable, too. The chain rule yields (4.1). Since $\mathscr{F}$ and $\mathscr{F}^{\prime}$ map bounded sets into bounded sets, by Theorem 4.3 , and since $\mathscr{L}$ is bounded, $\mathscr{E}$ and $\mathscr{E}^{\prime}$ map bounded sets into bounded sets, too.

We consider next the Banach space $V=W_{0}^{1, p}(\Omega) \cap L^{2}(\Omega)$, equipped with the sum norm $\|u\|_{V}:=\|u\|_{W^{1, p}}+\|u\|_{L^{2}}$, and the Hilbert space $H=L^{2}(\Omega)$, equipped with the usual inner product. Then $V$ is continuously and densely embedded into $L^{2}(\Omega)$ and into $W_{0}^{1, p}(\Omega)$. By Theorem 4.2, the restriction of $\mathscr{E}$ to the space $V$ is continuously differentiable, $\mathscr{E}$ and $\mathscr{E}^{\prime}$ map bounded sets into bounded sets, and

$$
\mathscr{E}^{\prime}(u) v=\int_{\Omega}|\nabla u|^{p-2} \nabla u \nabla v \quad \text { for every } u, v \in V
$$

The gradient of $\mathscr{E}$ with respect to the usual $L^{2}$ inner product is, according to the definition, the operator

$$
\begin{aligned}
& D\left(\nabla_{L^{2} \mathscr{E}}\right)=\left\{u \in V \text { : there exists } v \in L^{2}(\Omega)\right. \text { such that for every } \\
& \left.\qquad \varphi \in V \text { one has } \int_{\Omega}|\nabla u|^{p-2} \nabla u \nabla \varphi=\int_{\Omega} v \varphi\right\} \text { and } \\
& \nabla_{L^{2}} \mathscr{E}(u)=v .
\end{aligned}
$$

We write ${ }_{\Omega}^{D} \Delta_{p}:=-\nabla_{L^{2}} \mathscr{E}$ and we call this negative gradient the Dirichlet- $p$ Laplace operator on $L^{2}(\Omega)$. This term is justified by the following considerations.

The $p$-Laplace operator is the suitable partial differential operator which to every function $u: \Omega \rightarrow \mathbb{R}$ assigns the function

$$
\Delta_{p} u:=\operatorname{div}\left(|\nabla u|^{p-2} \nabla u\right)
$$

Note that $\Delta_{2}$ is equal to the Laplace operator $\Delta$.
Consider the boundary value problem

$$
\left\{\begin{align*}
-\Delta_{p} u & =f \text { in } \Omega  \tag{4.2}\\
u & =0 \text { on } \partial \Omega
\end{align*}\right.
$$

where $f \in L^{2}(\Omega)$ is a given function. We call a function $u \in W_{0}^{1, p}(\Omega) \cap L^{2}(\Omega)$ a weak solution of this problem if for every $\varphi \in W_{0}^{1, p}(\Omega) \cap L^{2}(\Omega)$ one has

$$
\int_{\Omega}|\nabla u|^{p-2} \nabla u \nabla \varphi=\int_{\Omega} f \varphi .
$$

With this definition of weak solution, by the definition of the Dirichlet- $p$-Laplace operator, and by Theorem 4.2, a function $u$ is a weak solution of the above boundary value problem if and only if

$$
-{ }_{\Omega}^{D} \Delta_{p} u=f
$$

### 4.4 An auxiliary result

Theorem 4.3. Let $\Omega \subseteq \mathbb{R}^{d}$ be an open set, and let $1<p<\infty$. Let $F: \Omega \times \mathbb{R}^{m} \rightarrow \mathbb{R}$ ( $m \geq 1$ ) be a measurable function such that

$$
F(x, \cdot) \text { is continuously differentiable for every } x \in \Omega, \text { and }
$$

$$
\text { there exist } C_{1} \in L^{p^{\prime}}(\Omega)\left(p^{\prime}:=\frac{p}{p-1}\right), C_{2} \in L^{1}(\Omega), C \geq 0 \text { such that }
$$

$$
\left|\frac{\partial}{\partial u} F(x, u)\right| \leq C_{1}(x)+C|u|^{p-1} \text { for every } x \in \Omega, u \in \mathbb{R}^{m}, \text { and }
$$

$$
|F(x, 0)| \leq C_{2}(x) \text { for every } x \in \Omega
$$

where $\frac{\partial}{\partial u} F(x, u)$ is the derivative of the function $F(x, \cdot)$ at the point $u$. Then the function $\mathscr{F}: L^{p}\left(\Omega ; \mathbb{R}^{m}\right) \rightarrow \mathbb{R}$ given by

$$
\mathscr{F}(u)=\int_{\Omega} F(x, u(x)) d x, \quad u \in L^{p}\left(\Omega ; \mathbb{R}^{m}\right)
$$

is well defined, continuously differentiable, and

$$
\mathscr{F}^{\prime}(u) v=\int_{\Omega} \frac{\partial}{\partial u} F(x, u(x)) v(x) d x \quad \text { for every } u, v \in L^{p}\left(\Omega ; \mathbb{R}^{m}\right)
$$

Moreover, $\mathscr{F}$ and $\mathscr{F}^{\prime}$ map bounded sets into bounded sets.
Proof. We first prove that the function $\mathscr{F}$ is well defined. The two growth conditions on $F$ and the mean value theorem imply that for every $x \in \Omega$ and every $u \in \mathbb{R}^{m}$

$$
\begin{aligned}
|F(x, u)| & \leq|F(x, 0)|+|F(x, u)-F(x, 0)| \\
& \leq|F(x, 0)|+\sup _{\xi \in[0, u]}\left|\frac{\partial F}{\partial u}(x, \xi)\right||u| \\
& \leq C_{2}(x)+\left(C_{1}(x)+C \sup _{\xi \in[0, u]}|\xi|^{p-1}\right)|u| \\
& =C_{2}(x)+C_{1}(x)|u|+C|u|^{p} .
\end{aligned}
$$

For every $u \in L^{p}\left(\Omega ; \mathbb{R}^{m}\right)$, Hölder's inequality implies that

$$
\int_{\Omega}|F(x, u(x))| d x \leq\left\|C_{2}\right\|_{L^{1}}+\left\|C_{1}\right\|_{L^{p^{\prime}}}\|u\|_{L^{p}}+C\|u\|_{L^{p}}^{p}
$$

By this estimate, the function $\mathscr{F}$ is well defined, and we also see that it maps bounded sets into bounded sets.

Fix $u \in L^{p}\left(\Omega ; \mathbb{R}^{m}\right)$ and define the linear functional $u^{\prime} \in L^{p}\left(\Omega ; \mathbb{R}^{m}\right)^{\prime}$ by

$$
u^{\prime}(v)=\int_{\Omega} \frac{\partial F}{\partial u}(x, u(x)) v(x), \quad v \in L^{p}\left(\Omega ; \mathbb{R}^{m}\right)
$$

This linear functional $u^{\prime}$ is well defined since, by Hölder's inequality, for every $v \in$ $L^{p}\left(\Omega ; \mathbb{R}^{m}\right)$

$$
\begin{align*}
\int_{\Omega}\left|\frac{\partial F}{\partial u}(x, u(x)) v(x)\right| d x & \leq\left\|\frac{\partial F}{\partial u}(\cdot, u)\right\|_{L^{p^{\prime}}}\|v\|_{L^{p}} \\
& \leq\left(\left\|C_{1}\right\|_{L^{p^{\prime}}}+C\|u\|_{L^{p}}^{p-1}\right)\|v\|_{L^{p}} \tag{4.3}
\end{align*}
$$

We prove that $\mathscr{F}$ is differentiable and that $\mathscr{F}^{\prime}(u)=u^{\prime}$. For this, we have to prove that

$$
\lim _{\|h\|_{L^{p}} \rightarrow 0} \frac{\left|\mathscr{F}(u+h)-\mathscr{F}(u)-u^{\prime}(h)\right|}{\|h\|_{L^{p}}}=0 .
$$

For every $h \in L^{p}\left(\Omega ; \mathbb{R}^{m}\right)$, by Hölder's inequality,

$$
\begin{align*}
& \left|\mathscr{F}(u+h)-\mathscr{F}(u)-u^{\prime}(h)\right| \leq  \tag{4.4}\\
& \leq \int_{\Omega}\left|F(x, u(x)+h(x))-F(x, u(x))-\frac{\partial F}{\partial u}(x, u(x)) h(x)\right| d x \\
& =\int_{\Omega}\left|\frac{F(x, u(x)+h(x))-F(x, u(x))-\frac{\partial F}{\partial u}(x, u(x)) h(x)}{|h(x)|}\right||h(x)| d x \\
& \leq\left(\int_{\Omega}\left|\frac{F(x, u(x)+h(x))-F(x, u(x))-\frac{\partial F}{\partial u}(x, u(x)) h(x)}{|h(x)|}\right|^{p^{\prime}} d x\right)^{\frac{1}{p^{\prime}}}\|h\|_{L^{p}}
\end{align*}
$$

where we interpret the term under the integral as 0 if $h(x)=0$. Let now $\left(h_{n}\right) \subset$ $L^{p}\left(\Omega ; \mathbb{R}^{m}\right)$ be a sequence converging to 0 . From this sequence, we can extract a subsequence (which we denote for simplicity again $\left(h_{n}\right)$ ) and we can find a $g \in$ $L^{p}(\Omega)$ such that

$$
\begin{aligned}
& h_{n} \rightarrow 0 \text { almost everywhere on } \Omega \text { and } \\
& \left|h_{n}\right| \leq g \text { almost everywhere, for every } n .
\end{aligned}
$$

Since $F(x, \cdot)$ is differentiable for every $x \in \Omega$, we have for this subsequence
$\left|\frac{F\left(x, u(x)+h_{n}(x)\right)-F(x, u(x))-\frac{\partial F}{\partial u}(x, u(x)) h_{n}(x)}{\left|h_{n}(x)\right|}\right| \rightarrow 0$ almost everywhere on $\Omega$.
Moreover, using the growth assumptions on $\frac{\partial F}{\partial u}$, for every $n$ and almost every $x$ we have

$$
\begin{aligned}
&\left|\frac{F\left(x, u(x)+h_{n}(x)\right)-F(x, u(x))-\frac{\partial F}{\partial u}(x, u(x)) h_{n}(x)}{\left|h_{n}(x)\right|}\right| \leq \\
& \leq \sup _{\xi(x) \in\left[u(x), u(x)+h_{n}(x)\right]}\left|\frac{\partial F}{\partial u}(x, \xi(x))\right|+\left|\frac{\partial F}{\partial u}(x, u(x))\right|
\end{aligned}
$$

484 Gradients in infinite dimensional space: the Laplace operator and the p-Laplace operator

$$
\begin{aligned}
& \leq 2 C_{1}(x)+C \sup _{\xi(x) \in\left[u(x), u(x)+h_{n}(x)\right]}|\xi(x)|^{p-1}+C|u(x)|^{p-1} \\
& \leq 2 C_{1}(x)+C\left(|u(x)|+\left.\left|h_{n}(x)\right|\right|^{p-1}+C|u(x)|^{p-1}\right. \\
& \leq 2 C_{1}(x)+C\left(C_{p}+1\right)|u(x)|^{p-1}+C C_{p}\left|h_{n}(x)\right|^{p-1} \\
& \leq 2 C_{1}(x)+C\left(C_{p}+1\right)|u(x)|^{p-1}+C C_{p} g(x)^{p-1} .
\end{aligned}
$$

Here, we have also used the estimate

$$
(a+b)^{p-1} \leq C_{p} a^{p-1}+C_{p} b^{p-1}
$$

which is true for every $a, b \geq 0$ and some constant $C_{p} \geq 0$. Hence, by Lebesgue's dominated convergence theorem,

$$
\lim _{n \rightarrow \infty} \int_{\Omega}\left|\frac{F\left(x, u(x)+h_{n}(x)\right)-F(x, u(x))-\frac{\partial F}{\partial u}(x, u(x)) h_{n}(x)}{\left|h_{n}(x)\right|}\right|^{p^{\prime}} d x=0 .
$$

We have thus proved that for every sequence $\left(h_{n}\right) \subset L^{p}\left(\Omega ; \mathbb{R}^{m}\right)$ converging to 0 there exists a subsequence (which we denote for simplicity again by $\left(h_{n}\right)$ ) such that the preceding equality holds. We claim that this implies

$$
\begin{equation*}
\lim _{\|h\|_{L^{p}} \rightarrow 0} \int_{\Omega}\left|\frac{F(x, u(x)+h(x))-F(x, u(x))-\frac{\partial F}{\partial u}(x, u(x)) h(x)}{|h(x)|}\right|^{p^{\prime}} d x=0 . \tag{4.5}
\end{equation*}
$$

In fact, if this was not true, then there exist $\varepsilon>0$ and a sequence $\left(h_{n}\right) \subset L^{p}\left(\Omega ; \mathbb{R}^{m}\right)$ converging to 0 such that

$$
\int_{\Omega}\left|\frac{F\left(x, u(x)+h_{n}(x)\right)-F(x, u(x))-\frac{\partial F}{\partial u}(x, u(x)) h_{n}(x)}{\left|h_{n}(x)\right|}\right|^{p^{\prime}} \geq \varepsilon,
$$

which is not possible by what we have proved above.
From (4.5) and the estimate (4.4) we deduce that $\mathscr{F}$ is differentiable and $\mathscr{F}^{\prime}(u)=$ $u^{\prime}$. In particular, $\mathscr{F}$ is continuous, and the estimate (4.3) says

$$
\left\|\mathscr{F}^{\prime}(u)\right\|_{\left(L^{p}\right)^{\prime}} \leq\left\|C_{1}\right\|_{L^{p^{\prime}}}+C\|u\|_{L^{p}}^{p-1},
$$

so that $\mathscr{F}^{\prime}$ maps bounded sets into bounded sets.
It remains only to prove that $\mathscr{F}^{\prime}$ is continuous. Fix $u \in L^{p}\left(\Omega ; \mathbb{R}^{m}\right)$, and let $\left(h_{n}\right) \subset$ $L^{p}\left(\Omega ; \mathbb{R}^{m}\right)$ be a sequence converging to 0 . As before, we can extract a subsequence (which we denote again by $\left(h_{n}\right)$ ) and we can find a $g \in L^{p}(\Omega)$ such that

$$
\begin{aligned}
& h_{n} \rightarrow 0 \text { almost everywhere on } \Omega \text { and } \\
& \left|h_{n}\right| \leq g \text { almost everywhere, for every } n .
\end{aligned}
$$

Then, using the growth assumptions on $\frac{\partial F}{\partial u}$ as above, by Hölder's inequality and by Lebesgue's dominated convergence theorem,

$$
\begin{aligned}
& \limsup _{n \rightarrow \infty}\left\|\mathscr{F}^{\prime}\left(u+h_{n}\right)-\mathscr{F}^{\prime}(u)\right\|_{\left(L^{p}\right)^{\prime}} \leq \\
& \leq \limsup _{n \rightarrow \infty} \sup _{\|v\|_{L^{p}} \leq 1} \int_{\Omega}\left|\frac{\partial F}{\partial u}\left(x, u(x)+h_{n}(x)\right)-\frac{\partial F}{\partial u}(x, u(x))\right||v(x)| d x \\
& \leq \limsup _{n \rightarrow \infty}\left(\int_{\Omega}\left|\frac{\partial F}{\partial u}\left(x, u(x)+h_{n}(x)\right)-\frac{\partial F}{\partial u}(x, u(x))\right|^{p^{\prime}} d x\right)^{\frac{1}{p^{\prime}}} \\
& =0
\end{aligned}
$$

In particular, limsup and liminf coincide. We have proved that for every sequence $\left(h_{n}\right) \subset L^{p}\left(\Omega ; \mathbb{R}^{m}\right)$ converging to 0 there exists a subsequence (which we denote again by $\left(h_{n}\right)$ ) such that

$$
\lim _{n \rightarrow \infty}\left\|\mathscr{F}^{\prime}\left(u+h_{n}\right)-\mathscr{F}^{\prime}(u)\right\|_{\left(L^{p}\right)^{\prime}}=0
$$

An argument by contradiction, as it was already used in this proof, implies that $\mathscr{F}^{\prime}$ is continuous.

### 4.5 Origins of the Laplace operator

... et j'ose me flatter de présenter aux géomètres, dans cet Ouvrage, une théorie des attractions des sphéroïdes et de la figure des planètes plus générale et plus simple que celles qui sont déjà connues. ${ }^{1}$

Pierre Simon de Laplace, 1785
In this short section, we try to describe the context in which the so-called Laplace operator and Laplace's equation appeared for the first time. As we already indicated, the Laplace operator arises in many partial differential equations describing various physical problems. One such physical problem is the problem of attraction of masses and Newton's gravitational force.

In 1785, in his memoir Théorie des attractions des sphérö̈des et de la figure des planètes ${ }^{2}$, Pierre Simon de Laplace describes the gravitational forces induced by mass distributions concentrated on spheroids (closed surfaces in the space), in particular by planets. It had been known that the gravitational force induced by a mass distribution is a vector field which is the gradient of a so-called potential.
... the recognition that a function exists which is a potential for the Newtonian gravitational force appeared first in Lagrange's 1773/1774 memoir Sur l'équation séculaire de la lune in

[^5]504 Gradients in infinite dimensional space: the Laplace operator and the p-Laplace operator
which it seems to serve principally as a convenient method for calculating the components of force in various coordinate systems. In 1777, Lagrange devoted a short paper to some of the properties of potential functions, including an application to the question of replacing a system of bodies by the centre of mass of the system for the purposes of calculating their motions. ${ }^{3}$

The gravitational potential, denoted by $V$ in Laplace's memoir, can be obtained by integrating the potential of a point mass over the mass distribution. The representation of the gravitational potential in this form allows one to consider very general mass distributions, in particular mass distributions on arbitrary spheroids, not only round spheres or other regular surfaces. This is an advantage compared to earlier works on celestial mechanics, as Laplace remarks in his introduction.

Considering spherical coordinates $(r, \varphi, \theta)$ and putting $\cos \theta=\mu$, Laplace obtains the equation ${ }^{4}$

$$
\frac{\partial}{\partial \mu}\left[\left(1-\mu^{2}\right) \frac{\partial V}{\partial \mu}\right]+\frac{1}{1-\mu^{2}} \frac{\partial^{2} V}{\partial \varphi^{2}}+r \frac{\partial^{2}(r V)}{\partial r^{2}}=0
$$

which is the so-called Laplace equation for the gravitational potential

$$
\Delta V=0
$$

in spherical coordinates.
In a later memoir from 1789 on Saturn rings ${ }^{5}$, Laplace writes the above equation in cartesian coordinates. He writes:
$V$ est la somme des molécules du sphérö̈de divisées par leurs distances respectives au point attiré $m$; pour avoir l'attraction du sphéroïde sur ce point parallèlement à une droite quelconque, il faut donc considérer $V$ comme une fonction de trois coordonnées rectangles dont l'une soit parallèle à cette droite, et différentier cette fonction relativement à cette coordonnée : le coefficient de la différentielle de la coordonnée, pris avec un signe contraire, sera la valeur de l'attraction du sphérö̈de décomposée parallèlement à la droite donnée et dirigée vers l'origine de la coordonnée qui lui est parallèle.

Si l'on représent par $\beta$ la fonction

$$
\frac{1}{\sqrt{\left(x-x^{\prime}\right)^{2}+\left(y-y^{\prime}\right)^{2}+\left(z-z^{\prime}\right)^{2}}},
$$

on aura

$$
V=\int \beta \rho d x^{\prime} d y^{\prime} d z^{\prime}
$$

[^6]... Mais il est facile de s'assurer, par la différentiation, que l'on a
$$
0=\frac{\partial^{2} \beta}{\partial x^{2}}+\frac{\partial^{2} \beta}{\partial y^{2}}+\frac{\partial^{2} \beta}{\partial z^{2}}
$$
on aura donc pareillement
$$
0=\frac{\partial^{2} V}{\partial x^{2}}+\frac{\partial^{2} V}{\partial y^{2}}+\frac{\partial^{2} V}{\partial z^{2}}
$$

Cette équation, rapportée à d'autres coordonnées, est la base de la théorie que j'ai présentée dans nos Mémoires de 1782 sur les attractions des sphérö̈des et sur la figure des planètes. ${ }^{6}$

The english translation is taken from [Jahnke (2003), pages 198-199]:
$V$ is the sum of the molecules of the spheroid divided by their respective distances to the attracted point. To find the attraction of the spheroid on this point parallel to a given line, one must therefore consider $V$ as a function of three rectangular coordinates of which one is parallel to this line, and differentiate this function with respect to this coordinate. The differential coefficient of the coordinate, taken with the opposite sign, is the value of the component of the attraction of the spheroid parallel to the given line and directed toward the origin of the coordinate which is parallel to it.

If we represent by $\beta$ the function

$$
\frac{1}{\sqrt{\left(x-x^{\prime}\right)^{2}+\left(y-y^{\prime}\right)^{2}+\left(z-z^{\prime}\right)^{2}}}
$$

[ $\beta$ is the graviational potential of a mass point] then we obtain

$$
V=\int \beta \rho d x^{\prime} d y^{\prime} d z^{\prime}
$$

[ $\rho$ is the density]... But it is easy to verify by differentiation that we have

$$
0=\frac{\partial^{2} \beta}{\partial x^{2}}+\frac{\partial^{2} \beta}{\partial y^{2}}+\frac{\partial^{2} \beta}{\partial z^{2}}
$$

likewise we have

$$
0=\frac{\partial^{2} V}{\partial x^{2}}+\frac{\partial^{2} V}{\partial y^{2}}+\frac{\partial^{2} V}{\partial z^{2}}
$$

Our translation of the last sentence of the french quotation is
This equation, given in different coordinates, is the basis of the theory which I presented in our Mémoires from $1782^{7}$ on the attraction of the spheroids and of the shape of the planets.

Laplace's memoirs from 1785 and 1789 are perhaps the first occasions on which Laplace's equation and the Laplace operator appear. The fact that graviational potentials $V$ satisfy Laplace's equation is a direct consequence of the fact that the potential $\beta$ satisfies Laplace's equation; this may be checked by a straightforward calculation. Note that the function $\beta$ is, up to a constant, the so-called fundamental

[^7]solution of Laplace's equation.

In the following, when studying the problem of heat conduction, we shall see how the Laplace operator appears as a consequence of Lemma 1.3, Fourier's law and the Divergence theorem (Theorem F.6). We shall see that the problem of heat conduction leads to Laplace's equation, too, if the heat conduction is stationary (constant in time); the unkown function $V$ then represents the heat density.

### 4.6 Exercises

4.1. Show that every function $u \in C_{c}^{1}(\Omega)\left(\Omega \subseteq \mathbb{R}^{d}\right.$ open) vanishes in a neighbourhood of the boundary $\partial \Omega$.
4.2. Let ${ }_{\Omega}^{D} \Delta$ be the Dirichlet-Laplace operator on $L^{2}(\Omega)$, and let $f \in L^{2}(\Omega)$. Show that $u \in H_{0}^{1}(\Omega)$ is a weak solution of the elliptic boundary value problem

$$
\left\{\begin{aligned}
-\Delta u & =f \text { in } \Omega \\
u & =0 \text { on } \partial \Omega
\end{aligned}\right.
$$

if and only if $u \in D\left({ }_{\Omega}^{D} \Delta\right)$ and

$$
-{ }_{\Omega}^{D} \Delta u=f .
$$

4.3 (The Neumann-Laplace operator on a bounded interval). Consider the Sobolev space $V=H^{1}(0,1)$ and the quadratic form $\mathscr{E}: H^{1}(0,1) \rightarrow \mathbb{R}$ given by

$$
\mathscr{E}(u)=\frac{1}{2} \int_{0}^{1}\left(u^{\prime}\right)^{2}
$$

Compute the gradient $\nabla_{L^{2}} \mathscr{E}$ with respect to the usual inner product in $H=L^{2}(0,1)$. More precisely, show that

$$
\begin{aligned}
D\left(\nabla_{L^{2}} \mathscr{E}\right) & =\left\{u \in H^{2}(0,1): u^{\prime}(0)=u^{\prime}(1)=0\right\} \\
\nabla_{L^{2}} \mathscr{E}(u) & =-u^{\prime \prime}
\end{aligned}
$$

Hint. One may use, without proof, the formula of integration by parts $\int_{0}^{1} v w^{\prime}=$ $v(1) w(1)-v(0) w(0)-\int_{0}^{1} v^{\prime} w$ for functions $v, w \in H^{1}(0,1)$. The inclusion $D\left(\nabla_{L^{2}} \mathscr{E}\right) \subset H^{2}(0,1)$ is shown similarly as in Section 3.4. In order to show that every $u \in D\left(\nabla_{L^{2}} \mathscr{E}\right)$ satisfies the boundary conditions $u^{\prime}(0)=u^{\prime}(1)=0$, use special "test functions" $\varphi \in H^{1}(0,1)$.
Remark. We call ${ }_{(0,1)}^{N} \Delta:=-\nabla_{L^{2}} \mathscr{E}$ the Neumann-Laplace operator on $L^{2}(0,1)$.
Solving the equation $-{ }_{(0,1)}^{N} \Delta u=f$ is equivalent to solving the boundary value problem

$$
\left\{\begin{array}{l}
u \in H^{2}(0,1)  \tag{4.6}\\
-u^{\prime \prime}(x)=f(x) \quad \text { for } x \in(0,1), \\
u^{\prime}(0)=u^{\prime}(1)=0
\end{array}\right.
$$

The boundary conditions $u^{\prime}(0)=0$ and $u^{\prime}(1)=0$ are called Neumann boundary conditions or no-flux boundary conditions.
4.4 (An Ornstein-Uhlenbeck operator with Dirichlet boundary conditions). Let $\Omega \subseteq \mathbb{R}^{d}$ be an open set, and let $B: \Omega \rightarrow \mathbb{R}$ be a bounded, continuously differentiable function with bounded gradient $\nabla B$. Consider the quadratic form $\mathscr{E}: H_{0}^{1}(\Omega) \rightarrow \mathbb{R}$ given by $\mathscr{E}(u)=\frac{1}{2} \int_{\Omega}|\nabla u|^{2} e^{-B}$, and let $\nabla_{H} \mathscr{E}$ be its gradient with respect to the inner product $\langle v, w\rangle_{H}=\int_{\Omega} v w e^{-B}$. Show that $H^{2}(\Omega) \cap H_{0}^{1}(\Omega) \subseteq D\left(\nabla_{H} \mathscr{E}\right)$, and that for every $u \in H^{2}(\Omega) \cap H_{0}^{1}(\Omega)$ one has $\nabla_{H} \mathscr{E}(u)=-\Delta u+\nabla B \nabla u$.
Hint. You may use, without proof, the product rule $\frac{\partial}{\partial x_{i}}(v w)=\frac{\partial v}{\partial x_{i}} w+v \frac{\partial w}{\partial x_{i}}$ for $v \in$ $H^{1}(\Omega), w \in C^{1}(\Omega)$.
4.5. Let $\mathscr{E}: V \rightarrow \mathbb{R}$ be a quadratic form on a Banach space $V$ which embeds densely and continuously into a Hilbert space $H$. Show that the gradient $\nabla_{H} \mathscr{E}$ is a symmetric operator on $H$, that is, for every $u, v \in D\left(\nabla_{H} \mathscr{E}\right)$ one has

$$
\left\langle\nabla_{H} \mathscr{E}(u), v\right\rangle_{H}=\left\langle u, \nabla_{H} \mathscr{E}(v)\right\rangle_{H}
$$

4.6. a) Let $F:(0,1) \times \mathbb{R} \rightarrow \mathbb{R}$ be a continuously differentiable function satisfying the hypotheses of Theorem 4.3 for some $p>1$, and let $f(x, u):=\frac{\partial F}{\partial u}(x, u)$. Show that the function $\mathscr{E}: H_{0}^{1}(0,1) \rightarrow \mathbb{R}$ given by

$$
\mathscr{E}(u)=\int_{0}^{1}\left[\frac{1}{2}\left(u^{\prime}\right)^{2}+F(\cdot, u)\right]
$$

is continously differentiable and compute the gradient with respect to the usual inner product in $L^{2}(0,1)$. More precisely, show that

$$
\begin{aligned}
D\left(\nabla_{L^{2}} \mathscr{E}\right) & =H^{2}(0,1) \cap H_{0}^{1}(0,1) \\
\nabla_{L^{2}} \mathscr{E}(u) & =-u^{\prime \prime}+f(\cdot, u)
\end{aligned}
$$

b) Let $F$ and $f$ be as in (a), and let further $m \in L^{\infty}(0,1)$ be a positive function such that $\frac{1}{m} \in L^{\infty}(0,1)$. Find a Banach space $V$, a function $\mathscr{E}: V \rightarrow \mathbb{R}$ and a Hilbert space $H$ such that $V \hookrightarrow H$ and

$$
\begin{aligned}
D\left(\nabla_{H} \mathscr{E}\right) & =H^{2}(0,1) \cap H_{0}^{1}(0,1) \\
\nabla_{H} \mathscr{E}(u) & =-m u^{\prime \prime}+f(\cdot, u)
\end{aligned}
$$

## 4.7 (The Dirichlet-Laplace operator).

a) (Poincaré's inequality). Show that $\langle u, v\rangle_{H_{0}^{1}}=\int_{0}^{1} u^{\prime} v^{\prime}$ defines an inner product on $H_{0}^{1}(0,1)$ which is equivalent to the usual $H^{1}$ inner product $\langle\cdot, \cdot\rangle_{H^{1}}$ (that

544 Gradients in infinite dimensional space: the Laplace operator and the p-Laplace operator
is, the corresponding norms are equivalent). Note that it suffices to prove that there exists a constant $\lambda>0$ such that

$$
\lambda \int_{0}^{1} u^{2} \leq \int_{0}^{1}\left(u^{\prime}\right)^{2} \quad \text { for every } u \in H_{0}^{1}(0,1)
$$

Hint. Prove this inequality first for $u \in C_{c}^{1}(0,1)$.
b) On the space $V=H^{2}(0,1) \cap H_{0}^{1}(0,1)$ we consider the quadratic form $\mathscr{E}: V \rightarrow$ $\mathbb{R}$ given by $\mathscr{E}(u)=\frac{1}{2} \int_{0}^{1}\left(u^{\prime \prime}\right)^{2}$. Compute the gradient of $\mathscr{E}$ with respect to the inner product $\langle\cdot, \cdot\rangle_{H_{0}^{1}}$ defined in (a). More precisely, show that

$$
\begin{aligned}
D\left(\nabla_{H_{0}^{1}} \mathscr{E}\right) & =\left\{u \in H^{3}(0,1) \cap H_{0}^{1}(0,1): u^{\prime \prime} \in H_{0}^{1}(0,1)\right\} \\
\nabla_{H_{0}^{1}} \mathscr{E}(u) & =-u^{\prime \prime}
\end{aligned}
$$

Hint. One may use the following two facts: first, for every $\psi \in C_{c}^{1}(0,1)$ there exists a (unique) solution of the problem

$$
\left\{\begin{array}{l}
\varphi \in C^{2}([0,1]) \\
-\varphi^{\prime \prime}(x)=\psi(x) \quad \text { for } x \in(0,1) \\
\varphi(0)=\varphi(1)=0
\end{array}\right.
$$

Note that $\varphi \in V$. Second, if $v, w \in L^{1}(0,1)$ and if $\int_{0}^{1} v \psi=\int_{0}^{1} w \psi$ for every $\psi \in$ $C_{c}^{1}(0,1)$, then $v=w$. The first fact may be proved by integrating the equation $-\varphi^{\prime \prime}=\psi$ twice and adjusting constants. For the second fact, see Lemma G.4.
4.8. a) Let $A: \mathbb{R}^{d} \rightarrow \mathbb{R}^{d}$ be a linear mapping and assume that $A=\nabla_{H} \mathscr{E}$ is the gradient of a quadratic form $\mathscr{E}: \mathbb{R}^{d} \rightarrow \mathbb{R}$ with respect to an inner product $\langle\cdot, \cdot\rangle_{H}$. Show that the spectrum of $A$ (considered also as a linear mapping $\mathbb{C}^{d} \rightarrow$ $\left.\mathbb{C}^{d}\right)$ is real, that is, whenever $A u=\lambda u$ for some $\lambda \in \mathbb{C}$ and some nonzero $u \in \mathbb{C}^{d}$, then $\lambda \in \mathbb{R}$.
b) (The damped mathematical pendulum). Let $\alpha, \beta>0$. Show that the ordinary differential equation of the damped mathematical pendulum

$$
\begin{array}{r}
\dot{u}_{1}-u_{2}=0 \\
\dot{u}_{2}+\alpha u_{2}+\beta u_{1}=0
\end{array}
$$

is not a gradient system of the form $\dot{u}+\nabla_{H} \mathscr{E}(u)=0$ if $\alpha<2 \sqrt{\beta}$ (that is, if the damping is weak).
Remark. By changing Exercise 2.3 in an appropriate way, one can show that the damped mathematical pendulum is a gradient system of the form $\dot{u}+\nabla_{g} \mathscr{E}(u)=0$, at least on $\mathbb{R}^{2} \backslash\{0\}$.

## Lecture 5

## Bochner-Lebesgue and Bochner-Sobolev spaces

This lecture on integration theory is a short introduction to the Bochner integral of Banach space-valued functions, to Bochner-Lebesgue spaces and to BochnerSobolev spaces. These notions are necessary notions in order to study abstract differential equations in Banach spaces and, in particular, gradient systems in Banach spaces; the solutions of such abstract differential equations naturally live in Bochner-Lebesgue or Bochner-Sobolev spaces of Banach space-valued functions.

We suppose that the reader is familiar with the Lebesgue integral of real-valued functions. While it is then relatively straightforward to define integrals, Lebesgue and Sobolev spaces of functions with values in $\mathbb{R}^{d}$ - by reducing everything to the components of such functions - the situation is a priori less clear if one considers functions with values in general Banach spaces. Should one argue componentwise, which means here, if $u: \Omega \rightarrow X$ is a function with values in a Banach space $X$, to reduce everything to the scalar-valued case by considering the functions $\left\langle x^{\prime}, u\right\rangle_{X^{\prime}, X}$ with $x^{\prime} \in X^{\prime}$ ? This seems to be one possibility, and if properly done, it leads to the definition of the so-called Pettis integral. For our purposes, however, we follow another strategy by repeating the ideas of the Lebesgue integral for functions with values in Banach spaces (notion of step function, measurability, integrability, integral). This leads to the definition of the Bochner integral which shares many properties of the Lebesgue integral, to Bochner-Lebesgue spaces and to Bochner-Sobolev spaces.

Here and throughout, we consider only open subsets in $\mathbb{R}^{d}$ with the Lebesgue measure as measure spaces, but most of the results on the Bochner integral and Bochner-Lebesgue spaces remain true for general measure spaces. We take the opportunity to prove some results for Bochner-Sobolev spaces of functions defined on intervals - which are equally true for functions with values in Banach spaces and for scalar-valued functions -, which were not proved in Lecture 3.

### 5.1 The Bochner integral

Let $X$ be a real Banach space with norm $\|\cdot\|$, let $\Omega \subseteq \mathbb{R}^{d}$ be an open set, and denote by $\mathscr{A}$ the Lebesgue $\sigma$-algebra on $\Omega$, that is, the smallest $\sigma$-algebra which contains the Borel $\sigma$-algebra (generated by the open sets) and all subsets of sets of Lebesgue measure zero. The Lebesgue measure on $\Omega$ is denoted by $\mu$.

A function $f: \Omega \rightarrow X$ is called step function, if there exists a sequence $\left(A_{n}\right) \subseteq \mathscr{A}$ of mutually disjoint Lebesgue measurable sets and a sequence $\left(x_{n}\right) \subset X$ such that $f=\sum_{n} 1_{A_{n}} x_{n}$, where $1_{A}$ denotes the characteristic function of the set $A$. A function $f: \Omega \rightarrow X$ is mesurable, if there exists a sequence $\left(f_{n}\right)$ of step functions $f_{n}: \Omega \rightarrow$ $X$ such that $f_{n} \rightarrow f$ pointwise almost everywhere. Note that in the case $X=\mathbb{R}$, this definition of measurability is equivalent to the one which says that preimages of measurable sets should be measurable. With the definition of measurability, the proofs of the statements (a)-(e) in the following lemma are straightforward.

Lemma 5.1. Let $X$ and $Y$ be two real Banach spaces.
a) Every continuous function $f: \Omega \rightarrow X$ is measurable.
b) If $f: \Omega \rightarrow X$ is measurable, then $\|f\|: \Omega \rightarrow \mathbb{R}$ is measurable.
c) If $f: \Omega \rightarrow X$ is measurable and if $g: X \rightarrow Y$ is continuous, then the composite function $g \circ f: \Omega \rightarrow Y$ is measurable.
d) If $f: \Omega \rightarrow X$ and $g: \Omega \rightarrow \mathbb{R}$ are measurable, then the product $f g: \Omega \rightarrow X$ is measurable.
e) If $f: \Omega \rightarrow X$ and $g: \Omega \rightarrow X^{\prime}$ are measurable, then the product $\langle g, f\rangle_{X^{\prime}, X}$ : $\Omega \rightarrow \mathbb{R}$ is measurable.
f) If $\left(f_{n}\right)$ is a sequence of measurable functions $\Omega \rightarrow X$ such that $f_{n} \rightarrow f$ pointwise almost everywhere, then $f$ is measurable.

The statement (f) may be obtained as a consequence of the corresponding statement with $X=\mathbb{R}$ and the following theorem which due to Pettis; see [Hille and Phillips (1957), Theorem 3.5.3].

Theorem 5.2 (Pettis). A function $f: \Omega \rightarrow X$ is measurable if and only if $\left\langle x^{\prime}, f\right\rangle$ is measurable for every $x^{\prime} \in X^{\prime}$ (we say that $f$ is weakly measurable) and there exists a Lebesgue null set $N \in \mathscr{A}$ such that $f(\Omega \backslash N)$ is separable.

We say that a function $f: \Omega \rightarrow X$ is integrable if $f$ is measurable and $\int_{\Omega}\|f\|<$ $\infty$, that is, if $f$ is measurable and the positive function $\|f\|: \Omega \rightarrow \mathbb{R}$ is integrable in the usual Lebesgue sense. For every integrable step function $f: \Omega \rightarrow X, f=$ $\sum_{n} 1_{A_{n}} x_{n}$, we define the (Bochner) integral

$$
\int_{\Omega} f d \mu:=\sum_{n} \mu\left(A_{n}\right) x_{n} .
$$

The series $\sum_{n} \mu\left(A_{n}\right) x_{n}$ converges absolutely and the limit is independent of the representation of $f$, as one can easily show. Hence, the Bochner integral for integrable
step functions is well defined; the integral $\int_{\Omega} f d \mu$ is an element of $X$. For every integrable function $f: \Omega \rightarrow X$ we define the (Bochner) integral

$$
\int_{\Omega} f d \mu:=\lim _{n \rightarrow \infty} \int_{\Omega} f_{n} d \mu
$$

where $\left(f_{n}\right)$ is any sequence of step functions $\Omega \rightarrow X$ such that $\left\|f_{n}\right\| \leq\|f\|$ and $f_{n} \rightarrow$ $f$ pointwise almost everywhere. We note without proof that such a sequence $\left(f_{n}\right)$ always exists, that the step functions $f_{n}$ are integrable, that the limit of the integrals $\int_{\Omega} f_{n} d \mu$ exists and that the limit is independent of the choice of the sequence $\left(f_{n}\right)$. The Bochner integral is therefore a natural generalization of the Lebesgue integral of a scalar-valued function. The Bochner integral enjoys many properties known from the Lebesgue integral. For example, the triangle inequality

$$
\left\|\int_{\Omega} f d \mu\right\| \leq \int_{\Omega}\|f\| d \mu
$$

is true. Lebesgue's dominated convergence theorem holds true, too.
Theorem 5.3 (Lebesgue, dominated convergence). Let $\left(f_{n}\right)$ be a sequence of integrable functions $\Omega \rightarrow X$ and let $f: \Omega \rightarrow X$ be a function. Suppose that there exists an integrable function $g: \Omega \rightarrow \mathbb{R}$ such that $\left\|f_{n}\right\| \leq g$ for every $n$ and $f_{n} \rightarrow f$ pointwise almost everywhere. Then $f$ is integrable and

$$
\int_{\Omega} f d \mu=\lim _{n \rightarrow \infty} \int_{\Omega} f_{n} d \mu
$$

It is easy to prove that the Bochner integral is linear in the following sense.

## Lemma 5.4 (Linearity of the Bochner integral).

a) For every integrable $f, g: \Omega \rightarrow X$ the sum $f+g$ is integrable and

$$
\int_{\Omega}(f+g) d \mu=\int_{\Omega} f d \mu+\int_{\Omega} g d \mu
$$

b) For every integrable $f: \Omega \rightarrow X$ and every linear continuous $T: X \rightarrow Y$ the function $T f: \Omega \rightarrow Y$ is integrable and

$$
\int_{\Omega} T f d \mu=T \int_{\Omega} f d \mu
$$

We also use the following notation for the Bochner integral:

$$
\int_{\Omega} f \text { or } \int_{\Omega} f(t) d \mu(t)
$$

and if $\Omega=(a, b)$ is an interval in $\mathbb{R}$, then we write

$$
\int_{a}^{b} f \text { or } \int_{a}^{b} f(t) d \mu(t) \text { or } \int_{a}^{b} f(t) d t
$$

for the Bochner integral of an integrable function $f:(a, b) \rightarrow X$.

### 5.2 Bochner-Lebesgue spaces

For every measurable function $f: \Omega \rightarrow X$ and every $1 \leq p<\infty$ we put

$$
\|f\|_{L^{p}}:=\left(\int_{\Omega}\|f\|^{p} d \mu\right)^{1 / p}
$$

We also put

$$
\|f\|_{L^{\infty}}:=\inf \{C \geq 0: \mu(\{\|f\| \geq C\})=0\}
$$

For every $1 \leq p \leq \infty$ we then define

$$
\mathscr{L}^{p}(\Omega ; X):=\left\{f: \Omega \rightarrow X \text { measurable }:\|f\|_{L^{p}}<\infty\right\} .
$$

Similarly as in the scalar case, one can show that $\mathscr{L}^{p}(\Omega ; X)$ is a vector space and that $\|\cdot\|_{L^{p}}$ is a seminorm on $\mathscr{L}^{p}(\Omega ; X)$. If we let

$$
\begin{aligned}
N_{p} & :=\left\{f \in \mathscr{L}^{p}(\Omega ; X):\|f\|_{L^{p}}=0\right\} \\
& =\left\{f \in \mathscr{L}^{p}(\Omega ; X): f=0 \text { almost everywhere }\right\}
\end{aligned}
$$

then the quotient space

$$
L^{p}(\Omega ; X):=\mathscr{L}^{p}(\Omega ; X) / N_{p}:=\left\{f+N_{p}: f \in \mathscr{L}^{p}(\Omega ; X)\right\}
$$

becomes a Banach space for the norm

$$
\|[f]\|_{L^{p}}:=\|f\|_{L^{p}} \quad\left([f]=f+N_{p}\right)
$$

The norm of the equivalence class $[f]$ is well defined, that is, it is independent of the representative $f$ in this class. We call the spaces $L^{p}(\Omega ; X)$ Bochner-Lebesgue spaces. As in the scalar case, we identify functions $f \in \mathscr{L}^{p}(\Omega ; X)$ with their equivalence classes $[f] \in L^{p}(\Omega ; X)$, and we say that $L^{p}$ is a function space. In particular, we identify two functions if they are equal almost everywhere.

If $\Omega=(a, b)$ is an interval in $\mathbb{R}$ we simply write

$$
L^{p}(a, b ; X):=L^{p}((a, b) ; X) .
$$

For bounded $\Omega \subseteq \mathbb{R}^{d}$ and if $1 \leq q \leq p \leq \infty$ we have the inclusions

$$
C(\bar{\Omega} ; X) \subseteq L^{\infty}(\Omega ; X) \subseteq L^{p}(\Omega ; X) \subseteq L^{q}(\Omega ; X) \subseteq L^{1}(\Omega ; X)
$$

as can be easily proved by using Hölder's inequality. In particular, if $\Omega$ is bounded and $f$ is continuous on the closure $\bar{\Omega}$, then $f$ belongs to $L^{p}(\Omega ; X)$ for every $1 \leq p \leq$ $\infty$.

Let us collect some results (without proof) about Bochner-Lebesgue spaces which are used in following lectures.

Theorem 5.5. If $1 \leq p<\infty$ and if $X$ is separable, then $L^{p}(\Omega ; X)$ is separable. More precisely, whenever $\left(h_{n}\right) \subseteq L^{p}(\Omega)$ and $\left(x_{n}\right) \subseteq X$ are two dense sequences, then

$$
\mathscr{F}:=\left\{f: \Omega \rightarrow X: f=h_{n} x_{m} \text { for some } n, m \in \mathbb{N}\right\}
$$

is countable and total in $L^{p}(\Omega ; X)$, that is, span $\mathscr{F}$ is dense in $L^{p}(\Omega ; X)$.
Let $1 \leq p \leq \infty$ and let $p^{\prime}=\frac{p}{p-1}$ be the conjugate exponent (with usual interpretations if $p=1$ or $p=\infty$ ). By Hölder's inequality, for every $f \in L^{p}(\Omega ; X)$ and every $g \in L^{p^{\prime}}\left(\Omega ; X^{\prime}\right)$ the function $\langle f, g\rangle_{X, X^{\prime}}$ is integrable and

$$
\left|\int_{\Omega}\langle f, g\rangle_{X, X^{\prime}}\right| \leq\|f\|_{L^{p}(\Omega ; X)}\|g\|_{L^{p^{\prime}}\left(\Omega ; X^{\prime}\right)}
$$

Moreover,

$$
\|g\|_{L^{p^{\prime}}}=\sup _{\|f\|_{L^{p}} \leq 1}\left|\int_{\Omega}\langle f, g\rangle_{X, X^{\prime}}\right|
$$

Thus, similarly as in the scalar case, we may identify $L^{p^{\prime}}\left(\Omega ; X^{\prime}\right)$ with a closed subspace of $L^{p}(\Omega ; X)^{\prime}$. The linear mapping which maps every $g \in L^{p^{\prime}}\left(\Omega ; X^{\prime}\right)$ to the continuous linear form $f \mapsto \int_{\Omega}\langle f, g\rangle$ is, as in the scalar case, an isometry. However, and this is different to the scalar case, this mapping is in general not surjective, even if $p<\infty$. The usual identification $L^{p}(\Omega)^{\prime} \cong L^{p^{\prime}}(\Omega)$ (for $p<\infty$ ) is no longer true for general Bochner-Lebesgue spaces, that is, we do not have a full identification of the dual space of $L^{p}(\Omega ; X)$ in terms of the Bochner integral and Bochner-Lebesgue spaces. However, the following results are true; for the proof of the first two statements, see [Diestel and Uhl (1977)], the third statement is an easy exercise.

Theorem 5.6. a) If $1 \leq p<\infty$ and if $X$ is reflexive, then $L^{p}(\Omega ; X)^{\prime} \cong L^{p^{\prime}}\left(\Omega ; X^{\prime}\right)$.
b) If $1<p<\infty$ and if $X$ is reflexive, then the space $L^{p}(\Omega ; X)$ is reflexive, too.
c) If $H$ is a Hilbert space, then the space $L^{2}(\Omega ; H)$ is a Hilbert space for the inner product

$$
\langle f, g\rangle_{L^{2}(\Omega ; H)}:=\int_{\Omega}\langle f, g\rangle_{H}, \quad f, g \in L^{2}(\Omega ; H)
$$

### 5.3 Bochner-Sobolev spaces in one space dimension

Let $X$ be a real Banach space, $-\infty \leq a<b \leq \infty$, and $p \in[1, \infty]$. The first BochnerSobolev space is the space
$W^{1, p}(a, b ; X):=\left\{u \in L^{p}(a, b ; X):\right.$ there exists $v \in L^{p}(a, b ; X)$ such that for

$$
\text { every } \left.\varphi \in C_{c}^{1}(a, b) \text { one has } \int_{a}^{b} u \varphi^{\prime}=-\int_{a}^{b} v \varphi\right\}
$$

By Lemma G.4, the function $v$ is uniquely determined, if it exists. We write $u^{\prime}:=v$ and we call $u^{\prime}$ the weak derivative of $u$. The spaces $W^{1, p}(a, b ; X)$ are Banach spaces for the norms

$$
\begin{aligned}
\|u\|_{W^{1, p}} & :=\left(\|u\|_{L^{p}}^{p}+\left\|u^{\prime}\right\|_{L^{p}}^{p}\right)^{\frac{1}{p}} \quad \text { if } 1 \leq p<\infty, \text { and } \\
\|u\|_{W^{1, \infty}} & :=\sup \left\{\|u\|_{L^{\infty}},\left\|u^{\prime}\right\|_{L^{\infty}}\right\}
\end{aligned}
$$

The linear mapping

$$
T: W^{1, p}(a, b ; X) \rightarrow L^{p}(a, b ; X) \times L^{p}(a, b ; X), \quad u \mapsto\left(u, u^{\prime}\right)
$$

shows that $W^{1, p}(a, b ; X)$ is isomorphic to a closed subspace of $L^{p}(a, b ; X) \times$ $L^{p}(a, b ; X)$. From this and from Theorems 5.5 and 5.6 we immediately obtain the following properties of Bochner-Sobolev spaces.

Theorem 5.7. a) If $1 \leq p<\infty$ and if $X$ is separable, then $W^{1, p}(a, b ; X)$ is separable.
b) If $1<p<\infty$ and if $X$ is reflexive, then the space $W^{1, p}(a, b ; X)$ is reflexive, too.
c) If $H$ is a Hilbert space, then the space $H^{1}(a, b ; H):=W^{1,2}(a, b ; H)$ is a Hilbert space for the inner product

$$
\langle u, v\rangle_{H^{1}(a, b ; H)}:=\int_{a}^{b}\langle u, v\rangle_{H}+\int_{a}^{b}\left\langle u^{\prime}, v^{\prime}\right\rangle_{H}, \quad u, v \in H^{1}(a, b ; H)
$$

In the rest of this section, we prove the analog of Theorem 3.3 and a little bit more.
Lemma 5.8. Let $u \in W^{1, p}(a, b ; X)$ be such that $u^{\prime}=0$. Then $u$ is constant almost everywhere.
Proof. Choose $\psi \in C_{c}^{1}(a, b)$ such that $\int_{a}^{b} \psi=1$. Then, for every $\varphi \in C_{c}^{1}(a, b)$, the function $\varphi-\left(\int_{a}^{b} \varphi\right) \psi$ is the derivative of a function in $C_{c}^{1}(a, b)$ since $\int_{a}^{b}(\varphi-$ $\left.\left(\int_{a}^{b} \varphi\right) \psi\right)=0$. Hence, by assumption and by definition of the weak derivative,

$$
0=\int_{a}^{b} u\left(\varphi-\left(\int_{a}^{b} \varphi\right) \psi\right) \quad \text { for every } \varphi \in C_{c}^{1}(a, b)
$$

If we put $c:=\int_{a}^{b} u \psi \in X$, then the preceding equality becomes

$$
\int_{a}^{b}(u-c) \varphi=0 \quad \text { for every } \varphi \in C_{c}^{1}(a, b)
$$

By Lemma G.4, $u=c$ almost everywhere.

Lemma 5.9. Let $(a, b)$ be a bounded interval, $t_{0} \in[a, b], g \in L^{p}(a, b ; X)$, and set

$$
u(t):=\int_{t_{0}}^{t} g(s) d s \quad \text { for every } t \in[a, b]
$$

Then $u \in W^{1, p}(a, b ; X)$ and $u^{\prime}=g$.
Proof. Let $\varphi \in C_{c}^{1}(a, b)$. Then, by Fubini's theorem,

$$
\begin{aligned}
\int_{a}^{b} u \varphi^{\prime} & =\int_{a}^{b} \int_{t_{0}}^{t} g(s) d s \varphi^{\prime}(t) d t \\
& =\int_{a}^{t_{0}} \int_{t_{0}}^{t} g(s) d s \varphi^{\prime}(t) d t+\int_{t_{0}}^{b} \int_{t_{0}}^{t} g(s) d s \varphi^{\prime}(t) d t \\
& =-\int_{a}^{t_{0}} \int_{a}^{s} \varphi^{\prime}(t) d t g(s) d s+\int_{t_{0}}^{b} \int_{s}^{b} \varphi^{\prime}(t) d t g(s) d s \\
& =-\int_{a}^{t_{0}} \varphi(s) g(s) d s-\int_{t_{0}}^{b} \varphi(s) g(s) d s \\
& =-\int_{a}^{b} g \varphi
\end{aligned}
$$

Theorem 5.10. Let $(a, b)$ be a bounded interval and $u \in W^{1, p}(a, b ; X)$. Then there exists a continuous function $\tilde{u}:[a, b] \rightarrow X$, which coincides with $u$ almost everywhere and such that for every s, $t \in[a, b]$

$$
\tilde{u}(t)-\tilde{u}(s)=\int_{s}^{t} u^{\prime}(r) d r
$$

Proof. Fix $t_{0} \in(a, b)$ and set $v(t):=\int_{t_{0}}^{t} u^{\prime}(s) d s$ for every $t \in[a, b]$. Clearly, the function $v$ is continuous. By Lemma 5.9, $v \in W^{1, p}(a, b ; X)$ and $v^{\prime}=u^{\prime}$. By Lemma 5.8, $u-v=c$ almost everywhere for some constant $c \in X$. This proves that $u$ coincides almost everywhere with the continuous function $\tilde{u}=v+c$, and that

$$
\tilde{u}(t)-\tilde{u}(s)=v(t)-v(s)=\int_{s}^{t} u^{\prime}(r) d r
$$

Due to Theorem 5.10, we may identify every function $u \in W^{1, p}(a, b ; X)$ with its continuous representative $\tilde{u}$, and we simply say that every function in $W^{1, p}(a, b ; X)$ is continuous. We state this in the following form.

Theorem 5.11 (Sobolev embedding theorem). Let $(a, b)$ be a bounded interval. Then $W^{1, p}(a, b ; X)$ is contained in $C([a, b] ; X)$ and there exists a constant $C \geq 0$ such that

$$
\|u\|_{L^{\infty}} \leq C\|u\|_{W^{1, p}} \quad \text { for every } u \in W^{1, p}(a, b ; X)
$$

Proof. The fact that every function $u \in W^{1, p}(a, b ; X)$ is continuous on $[a, b]$ follows from Theorem 5.10. The boundedness of the identity mapping $W^{1, p}(a, b ; X) \rightarrow$
$C([a, b] ; X), u \mapsto u$ may be seen as a consequence of the closed graph theorem. In fact, if $u_{n} \rightarrow u$ in $W^{1, p}(a, b ; X)$ and $u_{n} \rightarrow v$ in $C([a, b] ; X)$, then, since both spaces are continuously embedded into $L^{p}(a, b ; X), u_{n} \rightarrow u$ and $u_{n} \rightarrow v$ in $L^{p}(a, b ; X)$. By uniqueness of the limit, this is only possible if $u=v$. Hence, the identity mapping is closed.

Theorem 5.12 (Product rule, integration by parts). Let $(a, b)$ be a bounded interval, fix $1 \leq p \leq \infty$, and let $u \in W^{1, p}(a, b ; X)$ and $v \in W^{1, p}(a, b)$.
a) (Product rule). The product uv belongs to $W^{1, p}(a, b ; X)$ and

$$
(u v)^{\prime}=u^{\prime} v+u v^{\prime}
$$

b) (Integration by parts).

$$
\int_{a}^{b} u^{\prime} v=u(b) v(b)-u(a) v(a)-\int_{a}^{b} u v^{\prime}
$$

For every $1 \leq p \leq \infty$ and every $k \geq 2$ we define inductively the $k$-th BochnerSobolev spaces

$$
W^{k, p}(a, b ; X):=\left\{u \in W^{1, p}(a, b ; X): u^{\prime} \in W^{k-1, p}(a, b ; X)\right\},
$$

which are Banach spaces for the norms

$$
\begin{aligned}
\|u\|_{W^{k, p}} & :=\left(\sum_{j=0}^{k}\left\|u^{(j)}\right\|_{L^{p}}^{p}\right)^{\frac{1}{p}} \quad \text { if } 1 \leq p<\infty, \text { and } \\
\|u\|_{W^{1, \infty}} & :=\sup \left\{\|u\|_{L^{\infty}},\left\|u^{\prime}\right\|_{L^{\infty}}, \ldots,\left\|u^{(k)}\right\|_{L^{\infty}}\right\}
\end{aligned}
$$

If $H$ is a Hilbert space, then $H^{k}(a, b ; H):=W^{k, 2}(a, b ; H)$ is a Hilbert space for the inner product

$$
\langle u, v\rangle_{H^{k}}:=\sum_{j=0}^{k}\left\langle u^{(j)}, v^{(j)}\right\rangle_{L^{2}} .
$$

Finally, we define

$$
W_{0}^{1, p}(a, b ; X):=\overline{C_{c}^{1}(a, b ; X)}{ }^{\|\cdot\|_{W^{1, p}}, \text {, }, ~}
$$

and we put $H_{0}^{1}(a, b ; H):=W_{0}^{1,2}(a, b ; H)$.
Theorem 5.13. Let $(a, b)$ be a bounded interval. A function $u \in W^{1, p}(a, b ; X)$ belongs to $W_{0}^{1, p}(a, b ; X)$ if and only if $u(a)=u(b)=0$.

Proof. The "only if" part is a consequence of the definition of $W_{0}^{1, p}(a, b ; X)$ and the Sobolev embedding theorem (Theorem 5.11). The less obvious "if" part is left without proof.

Theorem 5.14 (Poincaré's inequality). Let $(a, b)$ be a bounded interval and $1 \leq$ $p<\infty$. Then there exists a constant $\lambda>0$ such that

$$
\lambda \int_{a}^{b}\|u\|^{p} \leq \int_{a}^{b}\left\|u^{\prime}\right\|^{p} \quad \text { for every } u \in W_{0}^{1, p}(a, b ; X)
$$

Proof. Let $u \in C_{c}^{1}(a, b ; X)$. Then

$$
\begin{aligned}
\int_{a}^{b}\|u(t)\|^{p} d t & =\int_{a}^{b}\left\|\int_{a}^{t} u^{\prime}(s) d s\right\|^{p} d t \\
& \leq \int_{a}^{b}\left(\int_{a}^{t}\left\|u^{\prime}(s)\right\| d s\right)^{p} d t \\
& \leq \int_{a}^{b}(t-a)^{p-1} \int_{a}^{t}\left\|u^{\prime}(s)\right\|^{p} d s d t \\
& \leq(b-a)^{p-1} \int_{a}^{b} \int_{a}^{b}\left\|u^{\prime}(s)\right\|^{p} d s d t \\
& =(b-a)^{p} \int_{a}^{b}\left\|u^{\prime}(s)\right\|^{p} d s
\end{aligned}
$$

This calculation yields Poincaré's inequality for every $u \in C_{c}^{1}(a, b ; X)$ with $\lambda=$ $(b-a)^{-p}$. Since, by definition, $C_{c}^{1}(a, b ; X)$ is dense in $W^{1, p}(a, b ; X)$, the full claim may be obtained by an approximation argument.

### 5.4 Exercises

5.1. Prove Theorem 5.12.

Hint. Prove the statement first for $u \in W^{1, p}(a, b ; X)$ and $v \in C^{1}([a, b])$.
5.2. Let $X$ and $Y$ be two Banach spaces such that $Y$ is continuously embedded into $X$. Fix $p \in[1, \infty)$ and let

$$
\mathscr{W}=W^{1, p}(0,1 ; X) \cap L^{p}(0,1 ; Y)
$$

be equipped with the norm

$$
\|u\|_{\mathscr{W}}^{p}=\|u\|_{L^{p}(0,1 ; Y)}^{p}+\left\|u^{\prime}\right\|_{L^{p}(0,1 ; X)}^{p}
$$

so that $\mathscr{W}$ is a Banach space. We recall from Theorem 5.10 that every function $u \in \mathscr{W}$ is continuous with values in $X$.
a) Show that the space $\mathscr{W}_{0}:=\{u \in \mathscr{W}: u(0)=0\}$ is a closed subspace of $\mathscr{W}$.
b) Show that the trace space

$$
\operatorname{Tr}:=\{x \in X: \text { there exists } u \in \mathscr{W} \text { such that } u(0)=x\}
$$

equipped with the norm

$$
\|x\|_{T r}:=\inf \left\{\|u\|_{\mathscr{W}}: u \in \mathscr{W} \text { and } u(0)=x\right\}
$$

is a Banach space.
Hint. Find a natural bijection $\mathscr{W} / \mathscr{W}_{0} \rightarrow T r$.
c) Show that $Y \subseteq \operatorname{Tr} \subseteq X$ and that the embeddings are continuous.
d) Show that if $X$ and $Y$ are Hilbert spaces and if $p=2$, then $\operatorname{Tr}$ is a Hilbert space.
e) Show that

$$
\mathscr{W} \subseteq C([0,1] ; \operatorname{Tr})
$$

with continuous embedding.
Remark. The trace space $\operatorname{Tr}$ is also denoted by $(X, Y)_{\frac{1}{p}, p}$ (with $p^{\prime}=\frac{p}{p-1}$ ). It is a particular example in the scale of so-called real interpolation spaces $(X, Y)_{\theta, p}$ which can be defined similarly by considering weighted $L^{p}$ spaces, [Lunardi (1995)]. In general, $\operatorname{Tr}$ is a strict subspace of $X$, so that the embedding from (e) is not a direct consequence of the Sobolev embedding theorem 5.11.
5.3. Let $\Omega \subseteq \mathbb{R}^{d}$ be open and bounded, and let $p \in[1, \infty)$.
a) Show that the mapping

$$
J: C([0,1] ; C(\bar{\Omega})) \rightarrow C([0,1] \times \bar{\Omega})
$$

given by $(J u)(t, x)=u(t)(x)((t, x) \in[0,1] \times \bar{\Omega})$ is bijective.
b) Show that for every $u \in C([0,1] ; C(\bar{\Omega}))$ one has

$$
\|u\|_{L^{p}\left(0,1 ; L^{p}(\Omega)\right)}=\|J u\|_{L^{p}((0,1) \times \Omega)} .
$$

c) Using the fact that $C_{c}(U)$ is dense in $L^{p}(U)$ whenever $U \subseteq \mathbb{R}^{m}$ is open and $p<\infty$ (Theorem G.3), show that $C([0,1] ; C(\bar{\Omega}))$ is dense in $L^{p}\left(0,1 ; L^{p}(\Omega)\right)$.
d) Conclude that $J$ extends to an isometric isomorphism $L^{p}\left(0,1 ; L^{p}(\Omega)\right) \rightarrow$ $L^{p}((0,1) \times \Omega)$.
Remark. While the above isomorphy seems to be natural, it is not immediately clear how to assign to a function (equivalence class) $u \in L^{p}\left(0,1 ; L^{p}(\Omega)\right)$ a function (equivalence class) $J u \in L^{p}((0,1) \times \Omega)$. The direct definition $(J u)(t, x)=u(t)(x)$ leads to the problem of measurability of $J u$. One way to circumvent this problem of measurability is the way described in this exercise.
5.4. Let $X$ and $Y$ be two Banach spaces such that $Y \hookrightarrow X$ and $X^{\prime} \hookrightarrow Y^{\prime}$ (dense and continuous embeddings), and let $p \in[1, \infty)$. Show that every function

$$
u \in W^{1, p}(0,1 ; X) \cap L^{\infty}(0,1 ; Y)
$$

(admits a representative which) is weakly continuous with values in $Y$, that is, for every $y^{\prime} \in Y^{\prime}$ the function $\left\langle y^{\prime}, u\right\rangle_{Y^{\prime}, Y}$ is continuous.

## Lecture 6 <br> Gradient systems in infinite dimensional spaces: existence and uniqueness of solutions

In this lecture we consider gradient systems in infinite-dimensional spaces. Under the assumption that the underlying energy is a convex function and under certain other assumptions we obtain existence and uniqueness of solutions. The proof of this key theorem (the proof of existence of a solution) is based on the so-called Ritz / Ritz-Galerkin / Faedo-Galerkin method. We see three advantages of this method. First, it is constructive, and it is therefore particularly interesting from the point of view of numerical analysis. Second, the method is intuitive and the four important steps of the proof are easy to remember (we hope that the reader agrees with this opinion; we do not claim that the whole proof with all technical details is easy to remember). Finally, we think that the method is efficient since it leads directly to a maximal regularity result.

Throughout, let $V$ be a Banach space with norm $\|\cdot\|_{V}$. Let $U \subseteq V$ be an open subset, and let $\mathscr{E}: U \rightarrow \mathbb{R}$ be a continuously differentiable function. In addition, let $H$ be a Hilbert space with inner product $\langle\cdot, \cdot\rangle_{H}$, and assume that $V$ is densely and continuously embedded into $H$. The gradient $\nabla_{H} \mathscr{E}: D\left(\nabla_{H} \mathscr{E}\right) \rightarrow H$ is defined as in Lecture 3.

### 6.1 Gradient systems in infinite dimensional spaces

A non-autonomous ${ }^{1}$ gradient system is a differential equation of the form

$$
\begin{equation*}
\dot{u}+\nabla_{H} \mathscr{E}(u)=f, \tag{6.1}
\end{equation*}
$$

[^8]where $\mathscr{E}$ as above and $f \in L_{l o c}^{2}(I ; H), I \subseteq \mathbb{R}$ being an interval (the subscript means that $f$ is square integrable on every compact subinterval of $I$ ). We call this gradient system non-autonomous since the function $f$ depends explicitly on the time variable.

A solution of the gradient system (6.1) is a measurable function $u: I \rightarrow V$ such that

$$
\begin{aligned}
& u \in W_{l o c}^{1,2}(I ; H) \cap L_{l o c}^{\infty}(I ; V) \\
& u(t) \in D\left(\nabla_{H} \mathscr{E}\right) \text { for almost every } t \in I, \text { and } \\
& \text { the equality (6.1) holds almost everywhere on } I .
\end{aligned}
$$

This definition of solution may look arbitrary and is motivated mainly by the existence and uniqueness result below, and by the energy inequality therein. We point out that the solutions in the above sense are exactly the solutions which have maximal possible regularity in the sense that if $u$ is a solution, then the two terms on the left-hand side of (6.1) have the same regularity (local square integrability) as the given right-hand side.

By the Sobolev embedding theorem (Theorem 5.11), every solution is continuous with values in $X$; this will allow us to give a sense to the initial value problem in Theorem 6.1. If $V=H$ is finite-dimensional and if $\nabla_{H} \mathscr{E}$ is continuous (that is, if $\mathscr{E}$ is continuously differentiable), then every solution in the above sense is also a solution as defined in Lecture 1, and in particular as considered in Carathéodory's theorem (Theorem 2.6). Conversely, if $u$ is a solution in the sense of Lecture 1, then $u$ is a solution in the above sense. This follows from the fact that the composite function $\nabla_{H} \mathscr{E}(u)$ is, by continuity, locally square-integrable, that for every $s, t \in I$

$$
u(t)-u(s)+\int_{s}^{t} \nabla_{H} \mathscr{E}(u(\tau)) d \tau=\int_{s}^{t} f(\tau) d \tau
$$

and from Lemma 5.9. Thus, in the finite-dimensional case, both notions of solutions coincide.

By definition of the gradient, $u$ is a solution of (6.1) if and only if $u \in W^{1,2}(I ; H) \cap$ $L^{\infty}(I ; V)$ and

$$
\begin{equation*}
\langle\dot{u}, v\rangle_{H}+\mathscr{E}^{\prime}(u) v=\langle f, v\rangle_{H} \quad \text { for every } v \in V \text { and almost every } t \in I \tag{6.2}
\end{equation*}
$$

We call (6.2) the variational form of the gradient system (6.1).

### 6.2 Global existence and uniqueness of solutions for gradient systems with convex energy

Let $\mathscr{E}: V \rightarrow \mathbb{R}$ be a function. We say that $\mathscr{E}$ is convex if
6.2 Global existence and uniqueness of solutions for gradient systems with convex energy

$$
\mathscr{E}(\lambda u+(1-\lambda) v) \leq \lambda \mathscr{E}(u)+(1-\lambda) \mathscr{E}(v) \quad \text { for every } u, v \in V, \lambda \in[0,1] .
$$

Moreover, we say that $\mathscr{E}$ is coercive if for every $c \in \mathbb{R}$ the sublevel set

$$
K_{c}=\{u \in V: \mathscr{E}(u) \leq c\} \text { is bounded in } V
$$

The following existence and uniqueness result for the gradient systems (6.1) with initial value is the main result of this lecture and a key result in the study of gradient systems. We say that it is a global existence result since the statement asserts existence of solutions on the whole interval on which the right-hand side $f$ is defined. Note that this interval is bounded in the statement, that is, $T<\infty$.

Theorem 6.1. Suppose that $V$ is a reflexive, separable Banach space, and suppose that $\mathscr{E}: V \rightarrow \mathbb{R}$ is a convex, coercive, continuously differentiable function and that $\mathscr{E}^{\prime}$ maps bounded sets into bounded sets. Then, for every $f \in L^{2}(0, T ; H)(T<\infty)$ and every $u_{0} \in V$ the gradient system with initial value

$$
\left\{\begin{array}{l}
\dot{u}+\nabla_{H} \mathscr{E}(u)=f  \tag{6.3}\\
u(0)=u_{0}
\end{array}\right.
$$

admits a unique solution $u \in W^{1,2}(0, T ; H) \cap L^{\infty}(0, T ; V)$. For this solution, and for every $t \in[0, T]$, the energy inequality

$$
\begin{equation*}
\int_{0}^{t}\|\dot{u}\|_{H}^{2}+\mathscr{E}(u(t)) \leq \mathscr{E}\left(u_{0}\right)+\int_{0}^{t}\langle f, \dot{u}\rangle_{H} \tag{6.4}
\end{equation*}
$$

holds true.

## Uniqueness

Uniqueness of solutions is a consequence of differentiability and convexity of $\mathscr{E}$, and the following lemma.

Lemma 6.2. Let $\mathscr{E}: V \rightarrow \mathbb{R}$ be a differentiable, convex function on a Banach space $V$. Then, for every $u, v \in V$,

$$
\left\langle\mathscr{E}^{\prime}(u)-\mathscr{E}^{\prime}(v), u-v\right\rangle_{V^{\prime}, V} \geq 0
$$

Proof. By convexity, for every $u, v \in V$ and every $\lambda \in(0,1)$

$$
\mathscr{E}(u+\lambda(v-u)) \leq(1-\lambda) \mathscr{E}(u)+\lambda \mathscr{E}(v)
$$

Hence,

$$
\frac{\mathscr{E}(u+\lambda(v-u))-\mathscr{E}(u)}{\lambda} \leq \mathscr{E}(v)-\mathscr{E}(u)
$$

Letting $\lambda \rightarrow 0$ in this inequality gives

$$
\left\langle\mathscr{E}^{\prime}(u), v-u\right\rangle_{V^{\prime}, V} \leq \mathscr{E}(v)-\mathscr{E}(u) .
$$

Changing the roles of $u$ and $v$ in this inequality gives

$$
\left\langle\mathscr{E}^{\prime}(v), u-v\right\rangle_{V^{\prime}, V} \leq \mathscr{E}(u)-\mathscr{E}(v) .
$$

Summing up the preceding two inequalities yields the claim.
Proof (Proof of Theorem 6.1-Uniqueness). Let $u_{1}$ and $u_{2}$ be two solutions of (6.3). Then, by Lemma 6.2, for almost every $t \in[0, T]$,

$$
\begin{aligned}
\frac{1}{2} \frac{d}{d t}\left\|u_{1}(t)-u_{2}(t)\right\|_{H}^{2} & =\left\langle u_{1}(t)-u_{2}(t), \dot{u}_{1}(t)-\dot{u}_{2}(t)\right\rangle_{H} \\
& =-\left\langle u_{1}(t)-u_{2}(t), \nabla_{H} \mathscr{E}\left(u_{1}(t)\right)-\nabla_{H} \mathscr{E}\left(u_{2}(t)\right)\right\rangle_{H} \\
& =-\left\langle u_{1}(t)-u_{2}(t), \mathscr{E}^{\prime}\left(u_{1}(t)\right)-\mathscr{E}^{\prime}\left(u_{2}(t)\right)\right\rangle_{V, V^{\prime}} \\
& \leq 0
\end{aligned}
$$

As a consequence, for almost every $t \in[0, T]$,

$$
\left\|u_{1}(t)-u_{2}(t)\right\|_{H}^{2} \leq\left\|u_{1}(0)-u_{2}(0)\right\|_{H}^{2}=0
$$

and therefore $u_{1}=u_{2}$.

## Existence

Existence of a solution of the gradient system (6.3) is proved by constructing a sequence of solutions of approximating gradient systems on finite-dimensional spaces, by extracting a convergent subsequence, and by showing that the limit is a solution we are looking for. For the existence part of the proof of Theorem 6.1, we need the following two compactness results from functional analysis; see Theorems E. 18 and D.57. We recall the necessary notions of weak convergence and weak* convergence.

Given a Banach space $X$, we say that a sequence $\left(u_{n}\right) \subseteq X$ converges weakly to an element $u \in X$ (and we write $u_{n} \rightharpoonup u$ ) if for every $u^{\prime} \in X^{\prime} \lim _{n \rightarrow \infty}\left\langle u^{\prime}, u_{n}\right\rangle_{X^{\prime}, X}=$ $\left\langle u^{\prime}, u\right\rangle_{X^{\prime}, X}$. If $X$ is a Hilbert space with inner product $\langle\cdot, \cdot\rangle_{X}$, then, by the RieszFréchet representation theorem (Theorem D.53), $u_{n} \rightharpoonup u$ if and only if for every $v \in X \lim _{n \rightarrow \infty}\left\langle u_{n}, v\right\rangle_{X}=\langle u, v\rangle_{X}$.

Let $X^{\prime}$ be the dual space of $X$. We say that a sequence $\left(u_{n}^{\prime}\right) \subseteq X^{\prime}$ converges weakly* to an element $u^{\prime} \in X^{\prime}$ (and we write $u_{n}^{\prime} \xrightarrow{\text { weak* }} u^{\prime}$ ) if for every $u \in X$ $\lim _{n \rightarrow \infty}\left\langle u_{n}^{\prime}, u\right\rangle_{X^{\prime}, X}=\left\langle u^{\prime}, u\right\rangle_{X^{\prime}, X}$.

Theorem 6.3. a) (Banach-Alaoglu). Let $X$ be a separable Banach space. Then every bounded sequence in $X^{\prime}$ admits a weak* convergent subsequence.
b) Every bounded sequence in a Hilbert space admits a weakly convergent subsequence.

Proof (Proof of Theorem 6.1 - Existence). To prove existence of a solution, we use the so called Ritz / Ritz-Galerkin / Faedo-Galerkin approximation. We consider a sequence of appropriate gradient systems on appropriate finite dimensional subspaces of $V$, for which local existence is known. Then, by proving appropriate norm and energy bounds, we show that the approximating problems admit global solutions. The bounds for the solutions are also used to extract weak and weak* convergent subsequences. Finally, we show that the weak limit is a solution of the gradient system (6.3). It turns out that especially in the crucial step of proving norm and energy bounds, the gradient structure of our problem is particularly helpful.

Part 1 (Formulation of approximating problems in finite-dimensional spaces): Let $\left(w_{n}\right)$ be an arbitrary sequence in $V$ such that $\operatorname{span}\left\{w_{n}: n \geq 1\right\}$ is dense in $V$; such a sequence exists because $V$ is separable.

For every $m \in \mathbb{N}$, we put

$$
V_{m}=\operatorname{span}\left\{w_{n}: 1 \leq n \leq m\right\},
$$

and we choose $u_{0}^{m} \in V_{m}$ such that

$$
\lim _{m \rightarrow \infty} u_{0}^{m}=u_{0} \quad \text { in } V
$$

This is possible since $\bigcup_{m} V_{m}$ is dense in $V$ and since $V_{m} \subseteq V_{m+1}$.
For every $m \in \mathbb{N}$, we consider the variational problem of finding $u_{m} \in$ $W_{l o c}^{1,2}\left(\left[0, T_{m}\right) ; V_{m}\right)$ such that

$$
\left\{\begin{array}{l}
\left\langle\dot{u}_{m}, v\right\rangle_{H}+\mathscr{E}^{\prime}\left(u_{m}\right) v=\langle f, v\rangle_{H}  \tag{6.5}\\
\quad \text { for every } v \in V_{m} \text { and almost every } t \in\left(0, T_{m}\right), \\
u_{m}(0)=u_{0}^{m}
\end{array}\right.
$$

Problem (6.5) is equivalent to the problem of finding a solution $u_{m} \in$ $W_{l o c}^{1,2}\left(\left[0, T_{m}\right) ; V_{m}\right)$ of the non-autonomous gradient system

$$
\left\{\begin{array}{l}
\dot{u}_{m}+\nabla_{H_{m}} \mathscr{E}_{m}\left(u_{m}\right)=P_{m} f  \tag{6.6}\\
u_{m}(0)=u_{0}^{m}
\end{array}\right.
$$

where $\mathscr{E}_{m}$ is the restriction of $\mathscr{E}$ to $V_{m}, H_{m}=V_{m}$ is equipped with the inner product induced by $H, \nabla_{H_{m}} \mathscr{E}_{m}$ is the gradient of $\mathscr{E}_{m}$ in $V_{m}$ with respect to the inner product $\langle\cdot, \cdot\rangle_{H}$, and $P_{m}: H \rightarrow H$ is the orthogonal projection from $H$ onto $V_{m}$ with respect to the inner product $\langle\cdot, \cdot\rangle_{H}$. Since $V_{m}$ is finite dimensional, for every $u \in V_{m}$ the gradient $\nabla_{H_{m}} \mathscr{E}_{m}(u)$ exists and belongs to $V_{m}$.

By the corollary to Carathéodory's theorem (Corollary 2.8), problem (6.6) admits a maximal solution $u_{m} \in W_{l o c}^{1,2}\left(\left[0, T_{m}\right) ; V_{m}\right)$ (remember that solutions in the sense of Carathéodory's theorem are weakly differentiable). Maximal solution
means that either $T_{m}=T$, or $T_{m}<T$ and the solution $u_{m}$ can not be extended to any larger interval. For every $m \in \mathbb{N}$, let $u_{m}$ be a maximal solution of (6.6).

Part 2 (Bounds for the solutions $u_{m}$ of the approximating problems): We show that the maximal solutions $u_{m}$ are global, that is, $T_{m}=T$, and that the sequence $\left(u_{m}\right)$ is bounded in appropriate function spaces. This part of the proof essentially repeats arguments from the proof of Theorem 2.10.

We multiply the equation (6.6) by $\dot{u}_{m}$ with respect to the inner product $\langle\cdot, \cdot\rangle_{H}$ (or: we take $v=\dot{u}_{m}$ in (6.5)), integrate the result over $[0, t]\left(t \in\left(0, T_{m}\right)\right)$, and apply the Cauchy-Schwarz inequality in order to obtain

$$
\begin{align*}
& \int_{0}^{t}\left\|\dot{u}_{m}(s)\right\|_{H}^{2} d s+\mathscr{E}\left(u_{m}(t)\right)-\mathscr{E}\left(u_{0}^{m}\right)=  \tag{6.7}\\
& =\int_{0}^{t}\left\langle f(s), \dot{u}_{m}(s)\right\rangle_{H} d s \\
& \leq \frac{1}{2} \int_{0}^{t}\|f(s)\|_{H}^{2} d s+\frac{1}{2} \int_{0}^{t}\left\|\dot{u}_{m}(s)\right\|_{H}^{2} d s
\end{align*}
$$

Since $\lim _{m \rightarrow \infty} u_{0}^{m}=u_{0}$ in $V$, and since $\mathscr{E}$ is continuous, we have $\lim _{m \rightarrow \infty} \mathscr{E}\left(u_{0}^{m}\right)=$ $\mathscr{E}\left(u_{0}\right)$. In particular, the sequence $\left(\mathscr{E}\left(u_{0}^{m}\right)\right)$ is bounded. Hence, there exists a constant $C \geq 0$ which is independent of $m$ such that, for every $t \in\left(0, T_{m}\right)$,

$$
\begin{equation*}
\frac{1}{2} \int_{0}^{t}\left\|\dot{u}_{m}(s)\right\|_{H}^{2} d s+\mathscr{E}\left(u_{m}(t)\right) \leq C+\frac{1}{2} \int_{0}^{T}\|f(s)\|_{H}^{2} d s \tag{6.8}
\end{equation*}
$$

The right-hand side of this inequality is independent of $m$ and $t \in\left(0, T_{m}\right)$, and the first term on the left-hand side is positive. This implies that the set $\left\{u_{m}(t): m \in\right.$ $\left.\mathbb{N}, t \in\left(0, T_{m}\right)\right\}$ is contained in the sublevel set $K_{c}$, where $c:=C+\frac{1}{2} \int_{0}^{T}\|f(s)\|_{H}^{2} d s$. By coercivity of $\mathscr{E}, K_{c}$ is bounded in $V$, and therefore

$$
\sup _{m \in \mathbb{N} t \in\left[0, T_{m}\right)} \sup _{n}\left\|u_{m}(t)\right\|_{V}<\infty
$$

Since the continuous, convex, coercive function $\mathscr{E}$ is bounded from below (see exercise!), we in addition deduce from the inequality (6.8) that

$$
\sup _{m \in \mathbb{N}}\left\|\dot{u}_{m}\right\|_{L^{2}\left(0, T_{m} ; H\right)}<\infty .
$$

Since $T_{m} \leq T$ is finite, this implies that for each $m \in \mathbb{N}$ the function $\dot{u}_{m}$ is integrable on $\left[0, T_{m}\right)$. Hence, $u_{m}$ extends to a continuous function on the closed interval $\left[0, T_{m}\right]$, and Carathéodory's theorem and the definition of maximal solution imply that this is only possible if $T_{m}=T$, that is, the $u_{m}$ are global.

From the preceding two inequalities and the continuous embedding $V \hookrightarrow H$ we obtain that

$$
\begin{equation*}
\left(u_{m}\right) \text { is bounded in } W^{1,2}(0, T ; H) \cap L^{\infty}(0, T ; V) . \tag{6.9}
\end{equation*}
$$

By assumption, the derivative $\mathscr{E}^{\prime}: V \rightarrow V^{\prime}$ maps bounded sets into bounded sets, so that the boundedness of $\left(u_{m}\right)$ in $L^{\infty}(0, T ; V)$ implies that

$$
\left(\mathscr{E}^{\prime}\left(u_{m}\right)\right) \text { is bounded in } L^{\infty}\left(0, T ; V^{\prime}\right)
$$

Part 3 (Extracting a convergent subsequence): The space $W^{1,2}(0, T ; H)$ is a Hilbert space and the spaces $L^{\infty}(0, T ; V) \cong L^{1}\left(0, T ; V^{\prime}\right)^{\prime}$ and $L^{\infty}\left(0, T ; V^{\prime}\right) \cong$ $L^{1}(0, T ; V)^{\prime}$ are dual spaces by reflexivity of $V$ (and $V^{\prime}$ ) and by Theorem 5.6. Moreover, the spaces $L^{1}\left(0,1 ; V^{\prime}\right)$ and $L^{1}(0,1 ; V)$ are separable by Theorem 5.5. Hence, by Theorem 6.3, there exist $u \in W^{1,2}(0, T ; H), v \in L^{\infty}(0, T ; V), \chi \in L^{\infty}\left(0, T ; V^{\prime}\right)$ and a subsequence of $\left(u_{m}\right)$ (which we denote for simplicity again by $\left(u_{m}\right)$ ) such that

$$
\begin{aligned}
& u_{m} \rightharpoonup u \text { in } W^{1,2}(0, T ; H), \\
& u_{m} \xrightarrow{\text { weak* }} v \text { in } L^{\infty}(0, T ; V), \text { and } \\
& \mathscr{E}^{\prime}\left(u_{m}\right) \xrightarrow{\text { weak* }} \chi \text { in } L^{\infty}\left(0, T ; V^{\prime}\right) .
\end{aligned}
$$

We leave it as an exercise to show that $u \in W^{1,2}(0, T ; H) \cap L^{\infty}(0, T ; V)$ and $u=v$. Moreover, it is an exercise to show that every continuous linear operator between two Hilbert spaces maps weakly convergent sequences into weakly convergent sequences. Since the operator $W^{1,2}(0, T ; H) \rightarrow L^{2}(0, T ; H), u \mapsto \dot{u}$ and the point evaluations $W^{1,2}(0, T ; H) \rightarrow H, u \mapsto u(0)$ and $W^{1,2}(0, T ; H) \rightarrow H, u \mapsto u(T)$ are linear and continuous (for the continuity of the point evalutions, use the Sobolev embedding theorem 5.11 in order to see this), the weak convergence of $\left(u_{m}\right)$ in $W^{1,2}(0, T ; H)$ therefore implies

$$
\begin{aligned}
& \dot{u}_{m} \rightharpoonup \dot{u} \text { in } L^{2}(0, T ; H), \\
& u_{m}(0) \rightharpoonup u(0) \text { in } H \text { and } \\
& u_{m}(T) \rightharpoonup u(T) \text { in } H .
\end{aligned}
$$

The above weak respectively weak* convergences in the respective function spaces mean that

$$
\begin{align*}
& \int_{0}^{T}\left\langle v, u_{m}\right\rangle_{V^{\prime}, V} \rightarrow \int_{0}^{T}\langle v, u\rangle_{V^{\prime}, V} \text { for every } v \in L^{1}\left(0, T ; V^{\prime}\right), \\
& \int_{0}^{T}\left\langle\dot{u}_{m}, v\right\rangle_{H} \rightarrow \int_{0}^{T}\langle\dot{u}, v\rangle_{H} \text { for every } v \in L^{2}(0, T ; H), \text { and }  \tag{6.10}\\
& \int_{0}^{T}\left\langle\mathscr{E}^{\prime}\left(u_{m}\right), v\right\rangle_{V^{\prime}, V} \rightarrow \int_{0}^{T}\langle\chi, v\rangle_{V^{\prime}, V} \text { for every } v \in L^{1}(0, T ; V) .
\end{align*}
$$

Part 4 (Showing that the limit $u$ is a solution): First of all, we have just seen that $u_{m}(0) \rightharpoonup u(0)$ in $H$. On the other hand, $u_{m}(0)=u_{0}^{m}$ by (6.5), and $u_{0}^{m} \rightarrow u_{0}$ in $V$ by the choice of the sequence $\left(u_{0}^{m}\right)$. Since $V$ is continuously embedded into $H$, we obtain $u(0)=u_{0}$, that is, $u$ satisfies the initial condition in (6.3). It remains to show that $u$ satisfies also the differential equation.

Let $w \in V_{m}$ and $\varphi \in L^{2}(0, T)$. For every $n \geq m$ we multiply equation (6.6) by $\varphi(\cdot) w$ with respect to the inner product $\langle\cdot, \cdot\rangle_{H}$, integrate over $(0, T)$ and obtain that

$$
\int_{0}^{T}\left\langle\dot{u}_{n}(t), \varphi(t) w\right\rangle_{H} d t+\int_{0}^{T}\left\langle\mathscr{E}^{\prime}\left(u_{n}(t)\right), \varphi(t) w\right\rangle_{V^{\prime}, V} d t=\int_{0}^{T}\langle f(t), \varphi(t) w\rangle_{H} d t
$$

Letting $n \rightarrow \infty$ in this last equality and using (6.10), we obtain

$$
\int_{0}^{T}\langle\dot{u}(t), \varphi(t) w\rangle_{H} d t+\int_{0}^{T}\langle\chi(t), \varphi(t) w\rangle_{V^{\prime}, V} d t=\int_{0}^{T}\langle f(t), \varphi(t) w\rangle_{H} d t
$$

Using the fact that $\left\{\varphi(\cdot) w: w \in \bigcup_{m} V_{m}, \varphi \in L^{2}(0, T)\right\}$ spans a dense subspace of $L^{2}(0, T ; V)$ (Theorem 5.5), we obtain for every $v \in L^{2}(0, T ; V)$

$$
\begin{equation*}
\int_{0}^{T}\langle\dot{u}, v\rangle_{H} d t+\int_{0}^{T}\langle\chi, v\rangle_{V^{\prime}, V} d t=\int_{0}^{T}\langle f, v\rangle_{H} d t \tag{6.11}
\end{equation*}
$$

It is left to show that $\chi=\mathscr{E}^{\prime}(u)$. We multiply equation (6.6) by $u_{m}$ with respect to the inner product $\langle\cdot, \cdot\rangle_{H}$, integrate the result over $(0, T)$, and obtain

$$
\begin{align*}
\int_{0}^{T} \mathscr{E}^{\prime}\left(u_{m}\right) u_{m} d t & =\int_{0}^{T}\left\langle f, u_{m}\right\rangle_{H} d t-\int_{0}^{T}\left\langle\dot{u}_{m}, u_{m}\right\rangle_{H} d t  \tag{6.12}\\
& =\int_{0}^{T}\left\langle f, u_{m}\right\rangle_{H} d t-\int_{0}^{T} \frac{1}{2} \frac{d}{d t}\left\|u_{m}\right\|_{H}^{2} d t \\
& =\int_{0}^{T}\left\langle f, u_{m}\right\rangle_{H} d t-\frac{1}{2}\left\|u_{m}(T)\right\|_{H}^{2}+\frac{1}{2}\left\|u_{0}^{m}\right\|_{H}^{2}
\end{align*}
$$

The weak convergence $u_{m}(T) \rightharpoonup u(T)$ in $H$ implies

$$
\|u(T)\|_{H}^{2} \leq \liminf _{m \rightarrow \infty}\left\|u_{m}(T)\right\|_{H}^{2}
$$

This estimate, the convergence $u_{0}^{m} \rightarrow u_{0}$ in $H$, the equality (6.12) and the weak convergence $u_{m} \rightharpoonup u$ in $L^{2}(0, T ; H)$ imply that

$$
\begin{align*}
\limsup _{m \rightarrow \infty} \int_{0}^{T} \mathscr{E}^{\prime}\left(u_{m}\right) u_{m} d t & \leq \int_{0}^{T}\langle f, u\rangle_{H} d t-\frac{1}{2}\|u(T)\|_{H}^{2}+\frac{1}{2}\left\|u_{0}\right\|_{H}^{2}  \tag{6.13}\\
& =\int_{0}^{T}\langle f, u\rangle_{H} d t-\int_{0}^{T} \frac{1}{2} \frac{d}{d t}\|u\|_{H}^{2} d t \\
& =\int_{0}^{T}\langle f, u\rangle_{H} d t-\int_{0}^{T}\langle\dot{u}, u\rangle_{H} d t \\
& =\int_{0}^{T}\langle\chi, u\rangle_{V^{\prime}, V} d t
\end{align*}
$$

In the last equality we have also used (6.11). By Lemma 6.2 and integration over $(0, T)$,

$$
\int_{0}^{T}\left\langle\mathscr{E}^{\prime}\left(u_{m}\right), u_{m}-u\right\rangle_{V^{\prime}, V} \geq \int_{0}^{T}\left\langle\mathscr{E}^{\prime}(u), u_{m}-u\right\rangle_{V^{\prime}, V}
$$

so that the weak ${ }^{*}$ convergences $u_{m} \xrightarrow{\text { weak* }} u$ and $\mathscr{E}^{\prime}\left(u_{m}\right) \xrightarrow{\text { weak* }} \chi$ imply

$$
\liminf _{m \rightarrow \infty} \int_{0}^{T}\left\langle\mathscr{E}^{\prime}\left(u_{m}\right), u_{m}\right\rangle_{V^{\prime}, V} \geq \int_{0}^{T}\langle\chi, u\rangle_{V^{\prime}, V}
$$

This inequality together with (6.13) implies

$$
\begin{equation*}
\lim _{m \rightarrow \infty} \int_{0}^{T}\left\langle\mathscr{E}^{\prime}\left(u_{m}\right), u_{m}\right\rangle_{V^{\prime}, V}=\int_{0}^{T}\langle\chi, u\rangle_{V^{\prime}, V} \tag{6.14}
\end{equation*}
$$

Let $v \in L^{\infty}(0, T ; V)$ be arbitrary, and set $w_{\lambda}=(1-\lambda) u+\lambda v$, where $\lambda \in(0,1)$. By Lemma 6.2 and integration over $(0, T)$,

$$
\int_{0}^{T}\left\langle\mathscr{E}^{\prime}\left(u_{m}\right)-\mathscr{E}^{\prime}\left(w_{\lambda}\right), u_{m}-w_{\lambda}\right\rangle_{V^{\prime}, V} \geq 0
$$

Using the definition of $w_{\lambda}$, this inequality can be rewritten as

$$
\begin{aligned}
& \lambda \int_{0}^{T}\left\langle\mathscr{E}^{\prime}\left(u_{m}\right), u-v\right\rangle_{V^{\prime}, V} \geq \\
& \geq \lambda \int_{0}^{T}\left\langle\mathscr{E}^{\prime}\left(w_{\lambda}\right), u-v\right\rangle_{V^{\prime}, V}+\int_{0}^{T}\left\langle\mathscr{E}^{\prime}\left(w_{\lambda}\right), u_{m}-u\right\rangle_{V^{\prime}, V}-\int_{0}^{T}\left\langle\mathscr{E}^{\prime}\left(u_{m}\right), u_{m}-u\right\rangle_{V^{\prime}, V}
\end{aligned}
$$

Letting $m \rightarrow \infty$ in this inequality, we obtain on using again the weak* convergences $u_{m} \xrightarrow{\text { weak* }} u, \mathscr{E}^{\prime}\left(u_{m}\right) \xrightarrow{\text { weak* }} \chi$, and (6.14) that

$$
\lambda \int_{0}^{T}\langle\chi, u-v\rangle_{V^{\prime}, V} \geq \lambda \int_{0}^{T}\left\langle\mathscr{E}^{\prime}\left(w_{\lambda}\right), u-v\right\rangle_{V^{\prime}, V}
$$

We divide by $\lambda>0$ and let $\lambda \rightarrow 0$, use the continuity of $\mathscr{E}^{\prime}$ and Lebesgue's dominated convergence theorem. Then

$$
\int_{0}^{T}\langle\chi, u-v\rangle_{V^{\prime}, V} \geq \int_{0}^{T}\left\langle\mathscr{E}^{\prime}(u), u-v\right\rangle_{V^{\prime}, V}
$$

Since $v \in L^{\infty}(0, T ; V)$ is arbitrary in this inequality, and since $u \in L^{\infty}(0, T ; V)$, we therefore obtain

$$
\int_{0}^{T}\langle\chi, v\rangle_{V^{\prime}, V} \geq \int_{0}^{T}\left\langle\mathscr{E}^{\prime}(u), v\right\rangle_{V^{\prime}, V} \quad \text { for every } v \in L^{\infty}(0, T ; V)
$$

This inequality implies

$$
\mathscr{E}^{\prime}(u)=\chi
$$

Replacing $\chi$ in (6.11) by $\mathscr{E}^{\prime}(u)$, it follows that $u$ is a solution of the differential equation in (6.2).

## Energy inequality

The following lemma is a consequence of the Hahn-Banach theorem and will not be proved here (see Corollary E.37).

Lemma 6.4. Let $X$ be a Banach space, $u_{n}, u \in X$, and let $\mathscr{F}: X \rightarrow \mathbb{R}$ be a continuous convex function. Then $u_{n} \rightharpoonup u$ implies $\mathscr{F}(u) \leq \liminf _{n \rightarrow \infty} \mathscr{F}\left(u_{n}\right)$.

Proof (Proof of Theorem 6.1 - Energy inequality). We maintain the notation from the preceding proof of existence of a solution. In particular, $\left(u_{m}\right)$ is the sequence of solutions of the approximating problems (6.6) which we obtained after extracting a subsequence, and $u$ is the solution of the gradient system (6.3) which we obtained as a weak limit of the sequence $\left(u_{m}\right)$.

We recall from (6.7) that for every $m$ one has the energy equality

$$
\begin{equation*}
\int_{0}^{t}\left\|\dot{u}_{m}(s)\right\|_{H}^{2} d s+\mathscr{E}\left(u_{m}(t)\right)=\mathscr{E}\left(u_{0}^{m}\right)+\int_{0}^{t}\left\langle f(s), \dot{u}_{m}(s)\right\rangle_{H} d s . \tag{6.15}
\end{equation*}
$$

We also recall that $\lim _{m \rightarrow \infty} u_{0}^{m}=u_{0}$ in $V$. The continuity of $\mathscr{E}$ thus yields

$$
\lim _{m \rightarrow \infty} \mathscr{E}\left(u_{0}^{m}\right)=\mathscr{E}\left(u_{0}\right) .
$$

The weak convergence $\dot{u}_{m} \rightharpoonup \dot{u}$ in $L^{2}(0, T ; H)$ implies that for every $t \in[0, T]$

$$
\lim _{m \rightarrow \infty} \int_{0}^{t}\left\langle f(s), \dot{u}_{m}(s)\right\rangle_{H} d s=\int_{0}^{t}\langle f(s), \dot{u}(s)\rangle_{H} d s
$$

By Lemma 6.4 applied with $X=L^{2}(0, T ; H)$ and $\mathscr{F}(v)=\int_{0}^{t}\|v\|_{H}^{2}$, the same weak convergence implies that for every $t \in[0, T]$

$$
\int_{0}^{t}\|\dot{u}(s)\|_{H}^{2} d s \leq \liminf _{m \rightarrow \infty} \int_{0}^{t}\left\|\dot{u}_{m}(s)\right\|_{H}^{2} d s .
$$

The weak convergence $u_{m} \rightharpoonup u$ in $W^{1,2}(0, T ; H)$ implies that for every $t \in[0, T]$ one has $u_{m}(t) \rightharpoonup u(t)$ in $H$. On the other hand, since $\left(u_{m}\right)$ is bounded in $L^{\infty}(0, T ; V)$, we deduce that for almost every $t \in[0, T]$ one has $u_{m}(t) \rightharpoonup u(t)$ in $V$. By Lemma 6.4 applied with $X=V$ and $\mathscr{F}=\mathscr{E}$, this implies that for almost every $t \in[0, T]$

$$
\mathscr{E}(u(t)) \leq \liminf _{m \rightarrow \infty} \mathscr{E}\left(u_{m}(t)\right) .
$$

Taking the limes inferior (as $m \rightarrow \infty$ ) on both sides of (6.15) yields the energy inequality (6.4). Theorem 6.1 is completely proved.

### 6.3 Exercises

6.1. Show that a quadratic form $\mathscr{E}: V \rightarrow \mathbb{R}$ on a Banach space $V$ is convex if and only if it is nonnegative, that is, $\mathscr{E} \geq 0$.
6.2. Let $\Omega \subset \mathbb{R}^{d}$ be open, and let $p \in(1, \infty)$. Show that the energy $\mathscr{E}: W^{1, p}(\Omega) \rightarrow \mathbb{R}$, $\mathscr{E}(u)=\frac{1}{p} \int_{\Omega}|\nabla u|^{p}$ is convex.
6.3. Let $V$ be a Banach space, and let $C \subseteq V$ be a bounded convex set.
a) Show that every continuous, convex function $\mathscr{E}: C \rightarrow \mathbb{R}$ is bounded from below.
b) Show that every continuous, convex, coercive function $\mathscr{E}: V \rightarrow \mathbb{R}$ is bounded from below.

Remark. In this exercise, continuity of $\mathscr{E}$ may be replaced by lower semicontinuity. The function $\mathscr{E}: C \rightarrow \mathbb{R}$ is lower semicontinuous if $\lim _{n \rightarrow \infty} u_{n}=u$ (with $u_{n}, u \in C$ ) implies $\mathscr{E}(u) \leq \liminf _{n \rightarrow \infty} \mathscr{E}\left(u_{n}\right)$.
There is a theorem which asserts that every continuous, convex, coercive function $\mathscr{E}: V \rightarrow \mathbb{R}$ on a reflexive Banach space attains its infimum (and in particular it is bounded from below; see Theorem E.38). While that result is nontrivial, this exercise here may be solved by using only the definition of convexity and continuity.
6.4 (De Giorgi - Energy dissipation rate). In this exercise, we come back to finite dimensional gradient systems. Let $\mathscr{E}: \mathbb{R}^{d} \rightarrow \mathbb{R}$ be a continuously differentiable function, and let $g: \mathbb{R}^{d} \rightarrow \operatorname{Inner}\left(\mathbb{R}^{d}\right)$ be a Riemannian metric. Show that a continuously differentiable function $u: I \rightarrow \mathbb{R}^{d}(I \subseteq \mathbb{R}$ an interval) is a solution of the gradient system

$$
\dot{u}+\nabla_{g} \mathscr{E}(u)=0
$$

if and only if, for every $t \in I$, the energy inequality

$$
\frac{d}{d t} \mathscr{E}(u(t)) \leq-\frac{1}{2}\|\dot{u}(t)\|_{g(u)}^{2}-\frac{1}{2}\left\|\nabla_{g} \mathscr{E}(u(t))\right\|_{g(u)}^{2}
$$

holds.
Remark. According to a remark in L. Ambrosio, Gradient flows in metric spaces and in the space of probability measures, and applications to Fokker-Planck equations with respect to log-concave measures, Bolletino della Unione Matematica Italiana IX 1 (2008), 223-240, this observation is due to E. De Giorgi; see E. De Giorgi, A. Marino and M. Tosques, Problems of evolution in metric spaces and maximal decreasing curves, Atti Accad. Naz. Lincei rend. Cl. Sci. Fis. Mat. Natur. 68 (1980), 180-187 and G. Perelman and A. Petrunin, Quasigeodesics and gradient curves in Alexandrov spaces, unpublished preprint (1994).

## Lecture 7 <br> Diffusion equations

As far as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality. ${ }^{1}$

Albert Einstein
Diffusion equations are prominent examples of gradient systems, and they provide us with the first examples of abstract gradient systems where we can apply the Theorem 6.1 on global existence and uniqueness of solutions.

Although it is not strictly necessary for this course on gradient systems, we spend some time on the derivation of the various mathematical models related to the problem of heat conduction, of diffusion of particles, and of denoising of images. While the assumptions which lead one to the linear heat equation - the heat conservation law and Fourier's law - may be criticized, it is certainly useful to keep them in mind throughout the following study. The comparison of the results obtained from the mathematical analysis of the various models and of the physical experiments can help to justify or reject a particular model. In this respect, the proof of existence and uniqueness of solutions is a test for the validity of a model. We consider two comparatively simple diffusion equations for which we can prove global existence and uniqueness of solutions.

### 7.1 Heat conduction

Let $\Omega \subseteq \mathbb{R}^{d}$ be an open set respresenting a spatial domain (so usually $d \leq 3$ ). We consider the problem of heat conduction inside $\Omega$. For this, we consider a function $u=u(t, x)$ which represents the heat density depending on the time variable $t \in[0, T]$ and on the space variable $x \in \Omega$. Given an open set $\mathscr{O} \subseteq \Omega$, the integral

[^9]$$
\int_{\mathscr{O}} u(t, x) d x
$$
is then the total amount of heat, at time $t$, inside the volume $\mathscr{O}$. This total amount of heat depends explicitly on the time variable; it changes infinitesimally only if heat is transported through the boundary $\partial \mathscr{O}$ or if there is a heat source or sink inside $\mathscr{O}$. This is expressed by the heat conservation law
\[

\underbrace{\frac{d}{d t} \int_{\mathscr{O}} u(t, x) d x}_{$$
\begin{array}{c}
\text { infinitesimal }  \tag{7.1}\\
\text { change of heat }
\end{array}
$$}=\underbrace{-\int_{\partial \mathscr{O}} \mathbf{j}(t, x) \mathbf{n}(x) d \sigma(x)}_{$$
\begin{array}{c}
\text { heat transport } \\
\text { through boundary }
\end{array}
$$}+\underbrace{\int_{\mathscr{O}} f(t, x) d x}_{$$
\begin{array}{c}
\text { heat production } \\
\text { or loss inside } \mathscr{O}
\end{array}
$$} .
\]

Here $\mathbf{j}=\mathbf{j}(t, x)$ is the heat flux vector which indicates into which direction and, by its length, how much heat is transported, $\mathbf{n}=\mathbf{n}(x)$ is the outer normal vector at $x \in \partial \mathscr{O}$, and the Euclidean inner product $\mathbf{j n}$ is the amount of heat which is transported into normal direction; heat, which is transported into tangential direction does not leave or enter the volume $\mathscr{O}$ and has therefore no influence on the infinitesimal change of total amount of heat. Note that the minus sign on the right-hand side is necessary, for if the heat flux points into $\mathscr{O}$, then the total amount of heat should increase and the left-hand side and the right-hand side should be positive. However, for inwards pointing $\mathbf{j}$, the inner product $\mathbf{j} \mathbf{n}$ is negative. The term $f=f(t, x)$ stands for a heat source or sink (depending on whether it is positive or negative), and it depends here on time and space. We have also assumed that $\mathscr{O}$ is of class $C^{1}$, and we denote by $\sigma$ the surface measure on $\partial \mathscr{O}$.

When we interchange the time derivative with the integral on the left-hand side (by assuming that the function $u$ is regular enough), and when we apply Gauß' divergence theorem (Theorem F.6) to the first term on the right-hand side, then equation (7.1) becomes

$$
\int_{\mathscr{O}} \frac{\partial}{\partial t} u(t, x) d x=-\int_{\mathscr{O}} \operatorname{div} \mathbf{j}(t, x) d x+\int_{\mathscr{O}} f(t, x) d x
$$

where $\operatorname{div} \mathbf{j}=\sum_{i=1}^{d} \frac{\partial \mathbf{j}_{i}}{\partial x_{i}}$ is the divergence of $\mathbf{j}$. The preceding formula is true for every volume $\mathscr{O} \subseteq \Omega\left(\mathscr{O}\right.$ of class $\left.C^{1}\right)$, which is only possible if

$$
\begin{equation*}
\frac{\partial}{\partial t} u(t, x)=-\operatorname{div} \mathbf{j}(t, x)+f(t, x) \quad \text { for every }(t, x) \in(0, T) \times \Omega \tag{7.2}
\end{equation*}
$$

Experiments show that heat is transported away from zones with higher heat density into zones with lower heat density. Recall from Lemma 1.3 that the negative euclidean gradient $-\nabla u$ (the gradient is to be taken in the $x$ variable only) points into the direction of steepest descent, that is, into the direction where $u$ decreases most rapidly. It is therefore natural to assume that the heat flux vector $\mathbf{j}$ and the negative temperature gradient $-\nabla u$ point into the same direction. In the simplest model, one
may assume that the length of the heat flux $\mathbf{j}$ and of the negative temperature gradient $-\nabla u$ are proportional to each other. This leads to the linear constitutive relation

$$
\begin{equation*}
\mathbf{j}(t, x)=-c \nabla u(t, x) \quad \text { (Fourier's law) } \tag{7.3}
\end{equation*}
$$

where $c>0$ is the heat conductivity characterizing the material; we assume here that $c$ does not depend on time or space, that is, the material filling $\Omega$ is ideally homogeneous. Inserting Fourier's law into (7.2) and noting that $\operatorname{div} \nabla=\Delta$ is the Laplace operator, we obtain the heat equation

$$
\begin{equation*}
\frac{\partial}{\partial t} u(t, x)=c \Delta u(t, x)+f(t, x) \quad \text { for every }(t, x) \in(0, T) \times \Omega \tag{7.4}
\end{equation*}
$$

If instead of Fourier's law (7.3) we consider a nonlinear constitutive relation with the heat conductivity depending on $u$ and $\nabla u, c=c(u, \nabla u)$, then we are lead to the nonlinear heat equation

$$
\begin{equation*}
\frac{\partial}{\partial t} u=\operatorname{div}(c(u, \nabla u) \nabla u)+f \quad \text { in }(0, T) \times \Omega \tag{7.5}
\end{equation*}
$$

The choice $c=c(|\nabla u|)=|\nabla u|^{p-2}$ leads to the nonlinear heat equation

$$
\frac{\partial}{\partial t} u=\Delta_{p} u+f \quad \text { in }(0, T) \times \Omega
$$

where $\Delta_{p}$ is the $p$-Laplace operator: $\Delta_{p} u=\operatorname{div}\left(|\nabla u|^{p-2} \nabla u\right)$. The choice $c=c(u)=$ $m u^{m-1}$ leads to the porous medium equation

$$
\frac{\partial}{\partial t} u=\Delta\left(u^{m}\right)+f \quad \text { in }(0, T) \times \Omega
$$

On the boundary $\partial \Omega$, we consider three types of behaviour. Either, we assume that the boundary is ideally conducting and that the heat density outside $\Omega$ is kept at a certain level,

$$
\left.u=u_{\Omega^{c}} \quad \text { on }(0, T) \times \partial \Omega \quad \text { (Dirichlet boundary condition }\right)
$$

or we assume that the boundary is ideally isolating and there is no heat flux through the boundary,

$$
\mathbf{j} \mathbf{n}=0 \quad \text { on }(0, T) \times \partial \Omega \quad \text { (Neumann boundary condition) },
$$

or we assume that the heat flux through the boundary depends on the difference between the heat density inside and outside $\Omega$, for example linearly,

$$
\mathbf{j n}=b\left(u-u_{\Omega^{c}}\right) \quad \text { on }(0, T) \times \partial \Omega \quad \text { (Robin boundary condition). }
$$

In the first and in the last equality, $u_{\Omega^{c}}$ is the heat density outside $\Omega$ and in the simple models one sets $u_{\Omega^{c}}=0$. Of course, one may imagine other types of boundary conditions, for example mixed boundary conditions (one part of the boundary $\partial \Omega$ might be conducting, while the other part is isolating).

In the case that the constitutive relation for the heat flux vector is Fourier's law, and if $u_{\Omega^{c}}=0$, the Dirichlet, Neumann and Robin boundary conditions can be rewritten as

$$
\begin{array}{ccc}
u=0 & \text { on }(0, T) \times \partial \Omega & \text { (Dirichlet boundary condition), } \\
\frac{\partial}{\partial \mathbf{n}} u=0 & \text { on }(0, T) \times \partial \Omega & \text { (Neumann boundary condition), }
\end{array}
$$

and

$$
\frac{\partial}{\partial \mathbf{n}} u+\frac{b}{c} u=0 \quad \text { on }(0, T) \times \partial \Omega \quad \text { (Robin boundary condition), }
$$

respectively. Here, $\frac{\partial}{\partial \mathbf{n}} u=\nabla u \mathbf{n}$ is the outer normal derivative of $u$.

### 7.2 The reaction-diffusion equation

Again, let $\Omega \subseteq \mathbb{R}^{d}$ be an open set representing a spatial domain. We consider now a function $u=u(t, x)$ which represents the concentration of a chemical component or the density of a biological population occupying the region $\Omega$. In this case, for a given open set $\mathscr{O} \subseteq \Omega$, the integral

$$
\int_{\mathscr{O}} u(t, x) d x
$$

is the total amount of the chemical component or of the population, at time $t$, inside the volume $\mathscr{O}$. The conservation law can be expressed similarly as in (7.1),

$$
\begin{equation*}
\frac{d}{d t} \int_{\mathscr{O}} u(t, x) d x=-\int_{\partial \mathscr{O}} \mathbf{j}(t, x) \mathbf{n}(x) d \sigma(x)+\int_{\mathscr{O}} f(t, x, u(t, x)) d x \tag{7.6}
\end{equation*}
$$

where $\mathbf{j}=\mathbf{j}(t, x)$ is now the drift vector of particles resp. individuals, and $f=$ $f(t, x, u)$ indicates the production of $u$ due to chemical reaction, resp. birth or death of individuals (depending on whether $f$ is positive or negative).

Experiments show that particles resp. individuals wander from regions where they are dense to regions where they are rarer, and that the drift rate is proportional to the norm of the gradient of the density and points to the opposite direction of the gradient. This is expressed by

$$
\begin{equation*}
\mathbf{j}(t, x)=-c \nabla u(t, x) \quad \text { (Fick's law). } \tag{7.7}
\end{equation*}
$$

Here, the diffusion coefficient $c$ might be constant, or depend on $u$ and/or on $\nabla u$. In this way we obtain the reaction-diffusion equation

$$
\begin{equation*}
\frac{\partial}{\partial t} u(t, x)=\operatorname{div}(c \nabla u)+f(\cdot, \cdot, u) \quad \text { in }(0, T) \times \Omega \tag{7.8}
\end{equation*}
$$

This equation is equipped with boundary conditions, too. The Dirichlet boundary conditions mean that there are no particles resp. individuals on the boundary (or that their number is fixed), while Neumann boundary conditions mean that particles resp. individuals can not leave the region $\Omega$.

### 7.3 The Perona-Malik model in image processing

Let $\Omega \subseteq \mathbb{R}^{d}$ be an open set ( $d=2$ or $d=3$ ). A black-and-white image is a function $u: \Omega \rightarrow \mathbb{R}$. At every point $x \in \Omega$, the value $u(x)$ stands for a grey-level. A reasonable image takes only values in the interval $[0,1]$; for example, the value 0 stands for a black pixel, the value 1 for a white pixel, and values in $(0,1)$ for a grey pixel with a grey-level between black and white (the role of 0 and 1 is just convention and may be inverted of course).

Suppose that $u_{0}$ describes the input image which is an image destroyed by a noise. The task is to remove the noise. One possible algorithm to denoise the image is to solve the problem

$$
\begin{cases}u_{t}-\operatorname{div}(c(|\nabla u|) \nabla u)=0 & \text { in }(0, T) \times \Omega,  \tag{7.9}\\ \frac{\partial}{\partial \mathbf{n}} u=0 & \text { on }(0, T) \times \partial \Omega, \\ u(0, x)=u_{0}(x) & \text { for } x \in \Omega,\end{cases}
$$

and to take $u(T,$.$) for some T>0$ as the output image. Perona and Malik [Perona and Malik (1990)] proposed to take the diffusion coefficient $c$ depending on the magnitude of the gradient, $|\nabla u|$, in such a way that if $|\nabla u|$ is small (in zones where the grey level does not vary too much), then the system (7.9) smoothes the image (it behaves like a diffusion equation), and if $|\nabla u|$ is large, then the diffusion (the regularization) is stopped and edges will be preserved (observe that $|\nabla u|$ is large where the image has edges). This is achieved by a function $c=c(|\nabla u|)$, where $c$ is decreasing, $c(0)=1$, and $\lim _{s \rightarrow \infty} c(s)=0$ (zero diffusion for large $s=|\nabla u|$ ). The functions $c(s)=(1+s)^{p-2}(p<2), c(s)=e^{-s}$ or $c(s)=e^{-s^{2}}$ are possible candidates, and according to the model it seems that the faster the function $c$ decreases as $s \rightarrow \infty$, the better the edges in the image are preserved.

### 7.4 Global existence and uniqueness of solutions

## The linear heat equation

We first consider the initial-boundary value problem

$$
\left\{\begin{array}{l}
u_{t}-u_{x x}=f \quad \text { in }(0, T) \times(0,1)  \tag{7.10}\\
u(t, 0)=u(t, 1)=0 \text { for } t \in(0, T) \\
u(0, x)=u_{0}(x) \quad \text { for } x \in(0,1)
\end{array}\right.
$$

where $u_{0}:[0,1] \rightarrow \mathbb{R}$ and $f:(0, T) \times(0,1) \rightarrow \mathbb{R}, f=f(t, x)$, are two given functions. This problem is a model for the heat conduction in a rod (one dimensional spatial domain), with ideally conducting end points, heat conductivity $c=1$, given initial heat distribution $u_{0}$ and heat source $f$.

We recall from Lecture 3 the observation that solving the time-independent / stationary problem

$$
\left\{\begin{array}{l}
u \in H^{2}(0,1) \\
-u_{x x}(x)=f(x) \text { for } x \in(0,1), \\
u(0)=u(1)=0
\end{array}\right.
$$

for a given $f \in L^{2}(0,1)$ is equivalent to solving the abstract equation

$$
-_{(0,1)}^{D} \Delta u=f,
$$

where the Dirichlet-Laplace operator ${ }_{(0,1)}^{D} \Delta u=-\nabla_{L^{2}} \mathscr{E}$ is the $L^{2}$-gradient of the energy $\mathscr{E}: H_{0}^{1}(0,1) \rightarrow \mathbb{R}$ given by

$$
\mathscr{E}(u)=\frac{1}{2} \int_{0}^{1}\left(u^{\prime}\right)^{2} .
$$

Without addressing the problem what could be a solution of the problem (7.10) (certainly it is a function of the two variables $t$ and $x$ ), it seems to be natural to replace the problem (7.10) by the abstract gradient system

$$
\left\{\begin{array}{l}
\dot{u}-{ }_{(0,1)}^{D} \Delta u=f,  \tag{7.11}\\
u(0)=u_{0},
\end{array}\right.
$$

in which the unknown $u$ is a function of the time variable only and takes its values in $L^{2}(0,1)$.

Theorem 7.1. For every $u_{0} \in H_{0}^{1}(0,1)$ and every $f \in L^{2}\left(0, T ; L^{2}(0,1)\right)(T>0)$ the gradient system (7.11) admits a unique solution $u \in H^{1}\left(0, T ; L^{2}(0,1)\right) \cap$ $L^{\infty}\left(0, T ; H_{0}^{1}(0,1)\right)$.

Proof. The domain of the energy $\mathscr{E}$, that is, the space $V=H_{0}^{1}(0,1)$ is separable and reflexive. The energy $\mathscr{E}$ is a continuous quadratic form, and therefore it is continuously differentiable by Proposition 3.2. The derivative $\mathscr{E}^{\prime}$ is linear and continuous, and therefore bounded. Moreover, since $\mathscr{E}$ is a nonnegative $(\mathscr{E} \geq 0), \mathscr{E}$ is convex by Exercise 6.1. By Poincaré's inequality (see Exercise 4.7 or Theorem 5.14), the inner product $\langle v, w\rangle_{H_{0}^{1}}=\int_{0}^{1} v^{\prime} w^{\prime}$ is equivalent to the usual inner product $\langle v, w\rangle_{H^{1}}=\int_{0}^{1} v w+\int_{0}^{1} v^{\prime} w^{\prime}$, and in particular there exists a constant $C>0$ such that

$$
\|v\|_{H^{1}}^{2} \leq C\|v\|_{H_{0}^{1}}^{2} \quad \text { for every } v \in H_{0}^{1}(0,1)
$$

in fact, one may take $C=\frac{1+\lambda}{\lambda}$, where $\lambda>0$ is the constant from Poincaré's inequality. As a consequence of this estimate, every sublevel of $\mathscr{E}$ is bounded in $H_{0}^{1}(0,1)$, that is, $\mathscr{E}$ is coercive. The claim now follows from Theorem 6.1.

## A nonlinear reaction-diffusion equation

We consider the nonlinear initial-boundary value problem

$$
\begin{cases}u_{t}-u_{x x}+u^{3}=f & \text { in }(0, T) \times(0,1),  \tag{7.12}\\ u_{x}(t, 0)=u_{x}(t, 1)=0 & \text { for } t \in(0, T), \\ u(0, x)=u_{0}(x) & \text { for } x \in(0,1),\end{cases}
$$

where $u_{0}:[0,1] \rightarrow \mathbb{R}$ and $f:(0, T) \times(0,1) \rightarrow \mathbb{R}, f=f(t, x)$, are two given functions. This is a model for the evolution of the concentration of a chemical component in an isolated channel (one dimensional spatial domain), undergoing diffusion with Fick's law and diffusion coefficient equal to 1 , reaction with production rate $-u^{3}$, and external supply $f$. Note that the "production" rate is really negative; compare the problem (7.12) with the reaction-diffusion equation (7.8) and note our convention to write the reaction term on a different side of the equality sign.

Similarly as in Exercise 4.3 on the Neumann-Laplace operator and in Exercise 4.6 , by using Theorem 4.3 , one shows that the energy $\mathscr{E}: H^{1}(0,1) \rightarrow \mathbb{R}$ given by

$$
\mathscr{E}(u)=\frac{1}{2} \int_{0}^{1}\left(u^{\prime}\right)^{2}+\frac{1}{4} \int_{0}^{1} u^{4}
$$

is continuously differentiable and that

$$
\begin{aligned}
D\left(\nabla_{L^{2}} \mathscr{E}\right) & =\left\{u \in H^{2}(0,1): u_{x}(0)=u_{x}(1)=0\right\} \\
\nabla_{L^{2}} \mathscr{E}(u) & =-u_{x x}+u^{3}
\end{aligned}
$$

It therefore seems to be natural to replace the problem (7.10) by the abstract gradient system

$$
\left\{\begin{array}{l}
\dot{u}+\nabla_{L^{2}} \mathscr{E}(u)=f  \tag{7.13}\\
u(0)=u_{0}
\end{array}\right.
$$

Theorem 7.2. For every $u_{0} \in H^{1}(0,1)$ and every $f \in L^{2}\left(0, T ; L^{2}(0,1)\right)$ the gradient system (7.13) admits a unique solution $u \in H^{1}\left(0, T ; L^{2}(0,1)\right) \cap L^{\infty}\left(0, T ; H^{1}(0,1)\right)$.

Proof. By Theorem 6.1, it suffices to check that the energy $\mathscr{E}$ is continuously differentiable, convex and coercive, and that $\mathscr{E}^{\prime}$ is maps bounded sets of $H^{1}(0,1)$ into bounded sets of $H^{1}(0,1)^{\prime}$ (we note that the energy space $H^{1}(0,1)$ is separable and reflexive). We have already observed that $\mathscr{E}$ is continuously differentiable, and we leave it as an exercise to show that $\mathscr{E}$ is convex and $\mathscr{E}^{\prime}$ bounded. By Hölder's inequality,

$$
\mathscr{E}(u) \geq \frac{1}{2} \int_{0}^{1}\left(u^{\prime}\right)^{2}+\frac{1}{4}\left(\int_{0}^{1} u^{2}\right)^{2}
$$

and this inequality implies that $\mathscr{E}$ is coercive.
Remark 7.3. The question of existence and uniqueness of local and/or global solutions becomes more difficult if we want to study the nonlinear diffusion equation

$$
\begin{cases}u_{t}-u_{x x}-u^{3}=f & \text { in }(0, T) \times(0,1)  \tag{7.14}\\ u_{x}(t, 0)=u_{x}(t, 1)=0 \text { for } t \in(0, T) \\ u(0, x)=u_{0}(x) & \text { for } x \in(0,1)\end{cases}
$$

in which - with respect to the problem (7.12) - only the sign of the nonlinearity changed. In order to guess the difference, it is instructive to compare the solutions of the two ordinary differential equations

$$
\dot{u}+u^{3}=0 \quad \text { and } \quad \dot{u}-u^{3}=0 .
$$

The second equation does not admit positive global solutions.

### 7.5 Exercises

7.1. Throughout, let $p>1$.
a) Show that the embedding $W^{1, p}(0,1) \hookrightarrow C([0,1])$ is compact. Conclude that the embedding $W^{1, p}(0,1) \hookrightarrow L^{p}(0,1)$ is compact, too.
Hint. Use the Arzelà-Ascoli theorem.
b) Show that the infimum

$$
\lambda_{1}:=\inf \left\{\frac{\int_{0}^{1}\left|u^{\prime}\right|^{p}}{\int_{0}^{1}|u|^{p}}: u \in W_{0}^{1, p}(0,1), u \neq 0\right\}
$$

is attained, that is, there exists $u \in W_{0}^{1, p}(0,1), u \neq 0$ such that $\lambda_{1} \int_{0}^{1}|u|^{p}=$ $\int_{0}^{1}\left|u^{\prime}\right|^{p}$. Recall from Poincaré's inequality that $\lambda_{1}>0$.
c) Let $f: \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function, $F(s):=\int_{0}^{s} f(r) d r$. Assume that there exist constants $c_{1}, c_{2}, c_{3} \geq 0$ and $\lambda<\lambda_{1}$ such that for every $s \in \mathbb{R}$

$$
|f(s)| \leq c_{1}+c_{2}|s|^{p-1} \text { and } F(s) \geq-c_{3}-\frac{\lambda}{p}|s|^{p}
$$

Show that the energy $\mathscr{E}: W_{0}^{1, p}(0,1) \rightarrow \mathbb{R}$ given by $\mathscr{E}(u)=\int_{0}^{1}\left[\frac{1}{p}\left|u^{\prime}\right|^{p}+F(u)\right]$ is continuously differentiable and coercive.
d) Show that for every $\lambda \in\left(0, \lambda_{1}\right]$ the energy $\mathscr{E}: H_{0}^{1}(0,1) \rightarrow \mathbb{R}, \mathscr{E}(u)=$ $\frac{1}{2} \int_{0}^{1}\left[\left|u^{\prime}\right|^{2}-\lambda|u|^{2}\right]$ is convex, and that it is not convex if $\lambda>\lambda_{1}$.
Remark. In (b), (c) and (d) one may replace the interval $(0,1)$ by an arbitrary open bounded set $\Omega \subseteq \mathbb{R}^{d}$. By the Rellich-Kondrachov theorem (see [Adams and Fournier (2003)], for every bounded $\Omega$ the embedding $W_{0}^{1, p}(\Omega) \hookrightarrow$ $L^{p}(\Omega)$ is compact.
7.2. Let $C: \mathbb{R}_{+} \rightarrow \mathbb{R}$ be a continuously differentiable function such that $C^{\prime}(0)=0$. Assume that there exist $p>1, c_{1} \geq 0, R \geq 0$ such that $C^{\prime}(s) \leq c_{1} s^{p-1}$ for every $s \geq R$. Let $\Omega \subseteq \mathbb{R}^{d}$ be an open and bounded.
a) Show that the energy

$$
\begin{aligned}
\mathscr{E}: W^{1, p}(\Omega) & \rightarrow \mathbb{R}, \\
u & \mapsto \int_{\Omega} C(|\nabla u|)
\end{aligned}
$$

is continuously differentiable, and that for every $u, h \in W^{1, p}(\Omega)$,

$$
\mathscr{E}^{\prime}(u) h=\int_{\Omega} C^{\prime}(|\nabla u|) \frac{\nabla u}{|\nabla u|} \nabla h,
$$

where the term under the integral is interpreted as 0 if $\nabla u=0$. Moreover, $\mathscr{E}^{\prime}: W^{1, p}(\Omega) \rightarrow W^{1, p}(\Omega)^{\prime}$ maps bounded sets into bounded sets.
b) Assume in addition that $C$ is increasing and convex. Show that the energy $\mathscr{E}$ from (a) is convex.
c) Show that the energy $\mathscr{E}$ is not coercive.
7.3. Let $\Omega \subset \mathbb{R}^{d}$ be open. We return to the model of heat conduction in $\Omega$, starting from the conservation law (7.2). Find an interpretation of the constitutive relation

$$
\mathbf{j}(t, x)=-c \nabla u(t, x)+b(x) u(t, x),
$$

in which $b: \Omega \rightarrow \mathbb{R}^{d}$ is a given function. Which partial differential equation does one obtain from this constitutive relation and (7.2)?

## Lecture 8

## Gradient systems in infinite dimensional spaces: existence and uniqueness of solutions II

The assumptions of Theorem 6.1 on the existence and uniqueness of global solutions of gradient systems are natural and easy to verify in applications arising, for example, from diffusion models; we refer to the problems (7.10) and (7.12) which were considered and solved in the previous lecture. However, there are examples of diffusion models / gradient systems to which Theorem 6.1 does not apply: this is the case for the problem (7.12) if the reaction term $u^{3}$ is dropped (the energy of the Laplace operator with Neumann boundary conditions is not coercive), for the problem (7.14), in which the reaction term had a different sign, and this is the case for the Perona-Malik model even if the diffusion coefficient $c$ is chosen appropriately so that the associated energy is convex; compare with Exercises 7.2 (c) and 8.3 (b). It is therefore necessary to think about more general existence and uniqueness theorems.

In this lecture we prove a variant of Theorem 6.1. We would like to call it a generalisation of Theorem 6.1, although it is not a strict generalisation. First, in Theorem 8.1 below, the energy is assumed to be $H$-elliptic instead of coercive and convex. This new property is weaker than coercivitiy and convexity. For example, the energy of the Neumann-Laplace operator and the energy associated with the Perona-Malik model are $L^{2}(\Omega)$-elliptic - if in the latter the diffusion coefficient is chosen appropriately. Second, we consider gradients taken with respect to general metrics. The price which we pay for both generalisations is that we can not prove uniqueness of solutions in general, and the embedding of the energy space $V$ into the Hilbert space $H$ is assumed to be compact. Using again the Ritz / Ritz-Galerkin / Faedo-Galerkin method in the proof, the compact embedding seems to be a necessary assumption. At the same time, it is an open problem whether the hypotheses of Theorem 8.1 imply uniqueness of solutions or not.

8 Gradient systems in infinite dimensional spaces: existence and uniqueness of solutions II

### 8.1 Global existence and uniqueness of solutions for gradient systems with elliptic energy

Let $V$ be a Banach space with norm $\|\cdot\|_{V}$. Let $U \subseteq V$ be an open subset, and let $\mathscr{E}: U \rightarrow \mathbb{R}$ be a continuously differentiable function. In addition, let $H$ be a Hilbert space with inner product $\langle\cdot, \cdot\rangle_{H}$, and assume that $V$ is densely and continuously embedded into $H$. Let finally $g: U \rightarrow \operatorname{Inner}(H)$ be a metric on $U$. A non-autonomous gradient system is a differential equation of the form

$$
\begin{equation*}
\dot{u}+\nabla_{g} \mathscr{E}(u)=f \tag{8.1}
\end{equation*}
$$

with $\mathscr{E}$ and $g$ as above and $f \in L_{l o c}^{2}(I ; H), I \subseteq \mathbb{R}$ being an interval. A solution of this gradient system is a measurable function $u: I \rightarrow V$ such that

$$
\begin{aligned}
& u \in W_{l o c}^{1,2}(I ; H) \cap L_{l o c}^{\infty}(I ; V) \\
& u(t) \in D\left(\nabla_{g} \mathscr{E}\right) \text { for almost every } t \in I, \text { and }
\end{aligned}
$$

the equality (8.1) holds almost everywhere on $I$.
By the Sobolev embedding theorem (Theorem 5.11), every solution is continuous with values in $H$; this allows us to give a sense to the initial value problem in Theorem 8.1. Moreover, by a similar reasoning as in Lecture 6, if $V=H$ is finitedimensional, then every solution in the above sense is also a solution as defined in Lecture 1 and vice versa. By definition of the gradient, $u$ is a solution of (8.1) if and only if $u \in W_{l o c}^{1,2}(I ; H) \cap L_{l o c}^{\infty}(I ; V)$ and

$$
\begin{equation*}
\langle\dot{u}, v\rangle_{g(u)}+\mathscr{E}^{\prime}(u) v=\langle f, v\rangle_{g(u)} \quad \text { for every } v \in V, \text { almost everywhere on } I . \tag{8.2}
\end{equation*}
$$

We call (8.2) the variational form of the gradient system (8.1).
We say that a function $\mathscr{E}: V \rightarrow \mathbb{R}$ is $H$-elliptic if there exists $\omega \in \mathbb{R}$ such that the function $\mathscr{E}_{\omega}: V \rightarrow \mathbb{R}, u \mapsto \mathscr{E}(u)+\frac{\omega}{2}\|u\|_{H}^{2}$ is coercive and convex. If the space $H$ is clear from the context, then we may simply say that $\mathscr{E}$ is elliptic. Note that if $\mathscr{E}_{\omega}$ is coercive (resp. convex) for some $\omega \in \mathbb{R}$, then $\mathscr{E}_{\omega^{\prime}}$ is coercive (resp. convex) for every $\omega^{\prime} \geq \omega$.

Theorem 8.1. Suppose that $V$ is a reflexive, separable Banach space which is compactly embedded into $H$. Suppose that $\mathscr{E}$ is an $H$-elliptic, continuously differentiable function such that $\mathscr{E}^{\prime \prime}: V \rightarrow V^{\prime}$ maps bounded sets into bounded sets. Let $T \in(0, \infty)$. Suppose in addition that the metric $g$ is continuous in the sense that

$$
\left.\begin{array}{l}
u_{n} \rightharpoonup u \text { in } W^{1,2}(0, T ; H), \\
u_{n} \xrightarrow{\text { weak* }} \text { u in } L^{\infty}(0, T ; V),  \tag{8.3}\\
v_{n} \rightharpoonup v \text { in } L^{2}(0, T ; H), \text { and } \\
w \in L^{2}(0, T ; H)
\end{array}\right\} \Rightarrow \int_{0}^{T}\left\langle v_{n}, w\right\rangle_{g\left(u_{n}\right)} \rightarrow \int_{0}^{T}\langle v, w\rangle_{g(u)} \text {. }
$$

Suppose finally that there exist two constants $c_{1}, c_{2}>0$ such that for every $u \in V$ and every $v \in H$

$$
c_{1}\|v\|_{H} \leq\|v\|_{g(u)} \leq c_{2}\|v\|_{H}
$$

Then, for every $f \in L^{2}(0, T ; H)$ and every initial value $u_{0} \in V$, there exists a solution $u \in W^{1,2}(0, T ; H) \cap L^{\infty}(0, T ; V)$ of the problem

$$
\left\{\begin{array}{l}
\dot{u}+\nabla_{g} \mathscr{E}(u)=f  \tag{8.4}\\
u(0)=u_{0}
\end{array}\right.
$$

If, in addition, the metric $g$ is constant - that is, $\langle\cdot, \cdot\rangle_{g(u)}=\langle\cdot, \cdot\rangle_{H}$ for every $u \in V$ -, then the above problem admits a unique solution. For this solution, the energy inequality

$$
\begin{equation*}
\int_{0}^{t}\|\dot{u}\|_{H}^{2}+\mathscr{E}(u(t)) \leq \mathscr{E}\left(u_{0}\right)+\int_{0}^{t}\langle f, \dot{u}\rangle_{H} \tag{8.5}
\end{equation*}
$$

holds true.
Remark 8.2. Theorem 8.1 is an $L^{2}$-maximal regularity result for the nonlinear problem (8.4) in the sense that for every $f \in L^{2}(0, T ; H)$ and every $u_{0} \in V$, the problem (8.4) admits a solution $u$ such that the two members $\dot{u}$ and $\nabla_{g} \mathscr{E}(u)$ of the left-hand side of (8.4) belong also to $L^{2}(0, T ; H)$.

## Existence

In the proof of Theorem 8.1, we need the following three lemmas.
Lemma 8.3 (Gronwall). Let $\varphi:[0, T] \rightarrow \mathbb{R}_{+}$be a nonnegative continuous function. Assume that there exist two constants $C, \omega \geq 0$ such that for every $t \in[0, T]$

$$
\varphi(t) \leq C+\omega \int_{0}^{t} \varphi(s) d s
$$

Then, for every $t \in[0, T]$,

$$
\varphi(t) \leq C e^{\omega t}
$$

Proof. Let $\varepsilon>0$ and set $\psi(t):=\varepsilon+C+\omega \int_{0}^{t} \varphi(s) d s(t \in[0, T])$. Then $\psi$ is a positive, nondecreasing, continuously differentiable function. By assumption and since $\varepsilon>0$, for every $t \in[0, T]$,

$$
\psi^{\prime}(t)=\omega \varphi(t) \leq \omega \psi(t)
$$

Dividing by $\psi(t)>0$, and integrating the resulting inequality yields that for every $t \in[0, T]$

$$
\psi(t) \leq \psi(0) e^{\omega t}=(C+\varepsilon) e^{\omega t}
$$

Since $\varphi(t) \leq \psi(t)$ by assumption, and since $\varepsilon>0$ is arbitrary, this implies the claim.

Lemma 8.4. a) (Chain rule). Let $X$ and $Y$ be two Banach spaces, $p \in[1, \infty)$, $(a, b) \subseteq \mathbb{R}$ a bounded interval, and $F: X \rightarrow Y$ a continuously differentiable function. Then, for every $u \in W^{1, p}(a, b ; X)$ the composite function $F \circ u$ belongs to $W^{1, p}(a, b ; Y)$ and $\frac{d}{d t}(F \circ u)(t)=F^{\prime}(u(t)) \dot{u}(t)$ for almost every $t$.
b) (Product rule). Let $H$ be a Hilbert space and $(a, b) \subseteq \mathbb{R}$ be an interval. Then, for every $u \in H^{1}(a, b ; H)$ the function $\|u\|_{H}^{2}$ belongs to $W^{1,1}(a, b)$ and $\frac{d}{d t}\|u(t)\|_{H}^{2}=2\langle u(t), \dot{u}(t)\rangle_{H}$ for almost every $t \in(a, b)$.

Proof (Sketch). (a) The claim is true for every $u \in C^{1}([a, b] ; X)$ by the classical chain rule. Since $C^{1}([a, b] ; X)$ is dense in $W^{1, p}(a, b ; X)$ (compare with Corollary G.13, where this claim is proved for $X=\mathbb{R}$; the proof of Corollary G. 13 remains valid in the case of vector-valued functions), the claim follows upon an approximation argument which uses also the Sobolev embedding theorem (Theorem 5.11).

The proof of (b) is very similar: the product rule is true for continuously differentiable functions $u \in H^{1}(a, b ; H)$, and the subspace of continuously differentiable functions is dense in $H^{1}(a, b ; H)$ by the vector-valued variant of Corollary G.13. The claim thus follows upon an approximation argument.

Lemma 8.5 (Aubin - Lions). Let $X$ and $Y$ be two Banach spaces such that $Y \hookrightarrow X$ with compact embedding. Then, for every $p \in(1, \infty)$, the embedding

$$
W^{1, p}(0, T ; X) \cap L^{\infty}(0, T ; Y) \hookrightarrow C([0, T] ; X)
$$

is compact, too.
Proof. Let $\left(u_{n}\right) \subseteq W^{1, p}(0, T ; X) \cap L^{\infty}(0, T ; Y)$ be a bounded sequence. It suffices to prove that $\left(u_{n}\right)$ is relatively compact in $C([0, T] ; X)$. For every $n \in \mathbb{N}$ and every $s$, $t \in[0, T], s \leq t$,

$$
\begin{aligned}
\left\|u_{n}(t)-u_{n}(s)\right\|_{X} & =\left\|\int_{s}^{t} \dot{u}_{n}(r) d r\right\|_{X} & & \text { (Theorem 5.10) } \\
& \leq \int_{s}^{t}\left\|\dot{u}_{n}(r)\right\|_{X} d r & & \text { (triangle inequality) } \\
& \leq\left\|\dot{u}_{n}\right\|_{L^{p}(0, T ; X)}(t-s)^{\frac{1}{p}} & & \text { (Hölder's inequality) } \\
& \leq C(t-s)^{\frac{1}{p^{\prime}}} & &
\end{aligned}
$$

where the constant $C \geq 0$ does not depend on $n$, and $p^{\prime}=\frac{p}{p-1}<\infty$. This inequality implies that the sequence $\left(u_{n}\right)$ is uniformly equicontinuous with values in $X$.

Since $\left(u_{n}\right)$ is uniformly bounded in $L^{\infty}(0, T ; Y)$, there exists a set $N$ of Lebesgue measure 0 such that for every $t \in[0, T] \backslash N$ the sequence $\left(u_{n}(t)\right)$ is bounded in $Y$. Since $Y$ is compactly embedded into $X$, we obtain that for every $t \in[0, T] \backslash N$ the sequence $\left(u_{n}(t)\right)$ is relatively compact in $X$. Since $[0, T] \backslash N$ is dense in $[0, T]$, the Arzelà-Ascoli theorem (Theorem B.43) implies that the sequence $\left(u_{n}\right)$ is relatively compact in $C([0, T] ; X)$.

Proof (Proof of Theorem 8.1 - Existence). We proceed similarly as in the proof of Theorem 6.1. In particular, we use the Ritz / Ritz-Galerkin / Faedo-Galerkin approximation.

Part 1 (Formulation of the finite dimensional approximating problems): We choose a total sequence $\left(w_{n}\right) \subseteq V$, we define the finite dimensional spaces $V_{m}$, and we choose $u_{0}^{m} \in V_{m}$ such that $u_{0}=\lim _{m \rightarrow \infty} u_{0}^{m}$, just as in Part 1 of the proof of Theorem 6.1.

For every $m \in \mathbb{N}$, we consider the variational problem of finding $u_{m} \in$ $W_{\text {loc }}^{1,2}\left([0, T) ; V_{m}\right)$ such that

$$
\left\{\begin{array}{l}
\left\langle\dot{u}_{m}, v\right\rangle_{g\left(u_{m}\right)}+\mathscr{E}^{\prime}\left(u_{m}\right) v=\langle f, v\rangle_{g\left(u_{m}\right)}  \tag{8.6}\\
\quad \text { for every } v \in V_{m}, \text { almost everywhere on }(0, T) \\
u_{m}(0)=u_{0}^{m}
\end{array}\right.
$$

Problem (8.6) is equivalent to the problem of finding a solution $u_{m} \in$ $W_{l o c}^{1,2}\left([0, T) ; V_{m}\right)$ of the non-autonomous gradient system in $V_{m}$

$$
\left\{\begin{array}{l}
\dot{u}_{m}+\nabla_{g_{m}} \mathscr{E}_{m}\left(u_{m}\right)=P_{m}^{u_{m}} f,  \tag{8.7}\\
u_{m}(0)=u_{0}^{m}
\end{array}\right.
$$

where $\mathscr{E}_{m}$ and $g_{m}$ are the restrictions of $\mathscr{E}$ and $g$, respectively, to $V_{m}, \nabla_{g_{m}} \mathscr{E}_{m}$ is the gradient of $\mathscr{E}_{m}$ in $V_{m}$ with respect to the metric $g_{m}$, and $P_{m}^{u_{m}}: H \rightarrow H$ is the orthogonal projection from $H$ onto $V_{m}$ with respect to the inner product $\langle\cdot, \cdot\rangle_{g\left(u_{m}\right)}$. Since $V_{m}$ is finite dimensional, for every $u \in V_{m}$ the gradient $\nabla_{g_{m}} \mathscr{E}_{m}(u)$ exists and belongs to $V_{m}$.

In order to obtain existence of a maximal solutions, we check that the function $F:(0, T) \times V_{m} \rightarrow V_{m},(t, u) \mapsto \nabla_{g_{m}} \mathscr{E}_{m}(u)-P_{m}^{u} f(t)$ satisfies the Carathéodory conditions. Since $g: V \rightarrow \operatorname{Inner}(H)$ is a metric, it maps norm convergent sequences in $V$ into strongly convergent sequences in $\operatorname{Inner}(H)$. Since $\operatorname{Inner}\left(V_{m}\right)$ is finitedimensional, norm convergence and strong convergence coincide, and therefore the restriction $g_{m}: V_{m} \rightarrow \operatorname{Inner}\left(V_{m}\right)$ is a Riemannian metric. By Lemma 2.4, the gradient $\nabla_{g_{m}} \mathscr{E}_{m}$ is continuous, and the first term in the definition of $F$ satisfies the Carathéodory conditions.

We check that the second term in the definition of $F$ satisfies the Carathéodory conditions, too. For every $u \in V_{m}$, let $Q_{m}(u) \in \mathscr{L}\left(V_{m}\right)$ be the operator given by $\left\langle Q_{m}(u) v, w\right\rangle_{H}=\langle v, w\rangle_{g_{m}(u)}=\langle v, w\rangle_{g(u)}\left(v, w \in V_{m}\right)$. Then, for every $u, v \in V_{m}$ and almost every $t \in[0, T]$,

$$
\left\langle Q_{m}(u) P_{m}^{u} f(t), v\right\rangle_{H}=\left\langle P_{m}^{u} f(t), v\right\rangle_{g(u)}=\left\langle f(t), P_{m}^{u} v\right\rangle_{g(u)}=\langle f(t), v\rangle_{g(u)}
$$

Since $g$ is a metric, the right-hand side depends continuously on $u \in V_{m}$. Since $V_{m}$ is finite-dimensional, weak convergence and norm convergence coincide, and therefore for almost every $t \in[0, T]$ the function $V_{m} \rightarrow V_{m}, u \mapsto Q_{m}(u) P_{m}^{u} f(t)$ is continuous. Since $g_{m}$ is continuous, $Q_{m}$ is continuous, too (compare with Lemma 2.3), and
since for every $u \in V_{m}$ the operator $Q_{m}(u)$ is invertible, we obtain that the second term in the definition of $F$ is continuous with respect to the second variable. Moreover, since $\langle\cdot, \cdot\rangle_{g(u)}$ is uniformly equivalent to $\langle\cdot, \cdot\rangle_{H}$, for almost every $t \in[0, T]$ and every $u \in V_{m}$,

$$
\left\|P_{m}^{u} f(t)\right\|_{H} \leq \frac{1}{c_{1}}\left\|P_{m}^{u} f(t)\right\|_{g(u)} \leq \frac{1}{c_{1}}\|f(t)\|_{g(u)} \leq \frac{c_{2}}{c_{1}}\|f(t)\|_{H} .
$$

Since $f \in L^{2}(0, T ; H)$, this implies that $F$ satisfies the Carathéodory conditions. By the corollary to Carathéodory's theorem (Corollary 2.8), problem (8.7) admits a maximal solution $u_{m} \in W_{l o c}^{1,2}\left(\left[0, T_{m}\right) ; V_{m}\right)$. Maximal means here that either $T_{m}=T$, or $T_{m}<T$ and the solution $u_{m}$ can not be extended to any larger interval. For every $m \in \mathbb{N}$, let $u_{m}$ be a maximal solution of (8.7).

Part 2 (Bounds for the solutions $u_{m}$ of the approximating problems): We show that the maximal solutions $u_{m}$ are global, that is, $T_{m}=T$, and that the sequence $\left(u_{m}\right)$ is bounded in appropriate function spaces. This part of the proof essentially repeats arguments from the proof of Theorem 2.10.

We multiply the equation (8.7) by $\dot{u}_{m}$ with respect to the inner product $\langle\cdot, \cdot\rangle_{g\left(u_{m}\right)}$ (or: we take $v=\dot{u}_{m}$ in (8.6)), integrate the result over $[0, t]\left(t \in\left(0, T_{m}\right)\right.$ ), and apply Lemma 8.4 (a) and the Cauchy-Schwarz inequality in order to obtain

$$
\begin{aligned}
& \int_{0}^{t}\left\|\dot{u}_{m}(s)\right\|_{g\left(u_{m}(s)\right)}^{2} d s+\mathscr{E}\left(u_{m}(t)\right)-\mathscr{E}\left(u_{0}^{m}\right)= \\
& =\int_{0}^{t}\left\langle f(s), \dot{u}_{m}(s)\right\rangle_{g\left(u_{m}(s)\right)} d s \\
& \leq \frac{1}{2} \int_{0}^{t}\|f(s)\|_{g\left(u_{m}(s)\right)}^{2} d s+\frac{1}{2} \int_{0}^{t}\left\|\dot{u}_{m}(s)\right\|_{g\left(u_{m}(s)\right)}^{2} d s .
\end{aligned}
$$

Since $\lim _{m \rightarrow \infty} u_{0}^{m}=u_{0}$ in $V$, and since $\mathscr{E}$ is continuous, we have $\lim _{m \rightarrow \infty} \mathscr{E}\left(u_{0}^{m}\right)=$ $\mathscr{E}\left(u_{0}\right)$. In particular, the sequence $\left(\mathscr{E}\left(u_{0}^{m}\right)\right)$ is bounded. Hence, there exists a constant $C_{1} \geq 0$ which is independent of $m$ and $t \in\left[0, T_{m}\right)$ such that

$$
\begin{equation*}
\frac{c_{1}}{2} \int_{0}^{t}\left\|\dot{u}_{m}(s)\right\|_{H}^{2} d s+\mathscr{E}\left(u_{m}(t)\right) \leq C_{1}+\frac{c_{2}}{2} \int_{0}^{T}\|f(s)\|_{H}^{2} d s \tag{8.8}
\end{equation*}
$$

In this estimate we have also used the assumption that the inner products $\langle\cdot, \cdot\rangle_{g\left(u_{m}\right)}$ are uniformly equivalent to the inner product $\langle\cdot, \cdot\rangle_{H}$. Let $\omega \in \mathbb{R}$ be such that $\mathscr{E}_{\omega}$ is convex and coercive. Then the preceding inequality can be rewritten as

$$
\begin{align*}
\frac{c_{1}}{2} \int_{0}^{t}\left\|\dot{u}_{m}(s)\right\|_{H}^{2} d s+\left\|u_{m}(t)\right\|_{H}^{2}+\mathscr{E}_{\omega}\left(u_{m}(t)\right) & \leq \\
& \leq C_{1}+\frac{\omega+2}{2}\left\|u_{m}(t)\right\|_{H}^{2}+\frac{c_{2}}{2} \int_{0}^{T}\|f(s)\|_{H}^{2} d s \tag{8.9}
\end{align*}
$$

We estimate the term $\frac{1}{2}\left\|u_{m}\right\|_{H}^{2}$ as follows:

$$
\begin{aligned}
\frac{1}{2}\left\|u_{m}(t)\right\|_{H}^{2} & =\frac{1}{2}\left\|u_{m}(0)\right\|_{H}^{2}+\int_{0}^{t} \frac{1}{2} \frac{d}{d s}\left\|u_{m}(s)\right\|_{H}^{2} d s \\
& =\frac{1}{2}\left\|u_{m}(0)\right\|_{H}^{2}+\int_{0}^{t}\left\langle\dot{u}_{m}(s), u_{m}(s)\right\rangle_{H} d s \\
& \leq C_{2}+\frac{c_{1}}{4(\omega+2)} \int_{0}^{t}\left\|\dot{u}_{m}(s)\right\|_{H}^{2} d s+\frac{\omega+2}{c_{1}} \int_{0}^{t}\left\|u_{m}\right\|_{H}^{2} d s
\end{aligned}
$$

where $C_{2} \geq 0$ is an upper bound for the sequence $\left(\frac{1}{2}\left\|u_{0}^{m}\right\|_{H}^{2}\right)$. Since $\mathscr{E}_{\omega}$ is continuous, convex and coercive, $\mathscr{E}_{\omega}$ is bounded from below by Exercise 6.3 (b), that is, there exists a constant $C_{3} \geq 0$ such that $\mathscr{E}_{\omega}(u) \geq-C_{3}$ for every $u \in V$. Inserting this estimate and the preceding estimate into (8.9), we obtain

$$
\begin{aligned}
& \frac{c_{1}}{4} \int_{0}^{t}\left\|\dot{u}_{m}(s)\right\|_{H}^{2} d s+\left\|u_{m}(t)\right\|_{H}^{2} \leq \\
& \leq C_{1}+(\omega+2) C_{2}+C_{3}+\frac{(\omega+2)^{2}}{c_{1}} \int_{0}^{t}\left\|u_{m}(s)\right\|_{H}^{2} d s+\frac{c_{2}}{2} \int_{0}^{T}\|f(s)\|_{H}^{2} d s \\
& =C_{4}+\frac{(\omega+2)^{2}}{c_{1}} \int_{0}^{t}\left\|u_{m}(s)\right\|_{H}^{2} d s
\end{aligned}
$$

where $C_{4} \geq 0$ does not depend on $m$ and $t \in\left[0, T_{m}\right)$. The first term on the left-hand side is positive. Hence, by applying Gronwall's inequality to the function $\varphi(t)=$ $\left\|u_{m}(t)\right\|_{H}^{2}$, we obtain that for every $m$ and every $t \in\left[0, T_{m}\right)$,

$$
\left\|u_{m}(t)\right\|_{H}^{2} \leq C_{4} e^{\frac{(\omega+2)^{2}}{c_{1}} t} \leq C_{4} e^{\frac{(\omega+2)^{2}}{c_{1}} T}
$$

The right-hand side of this inequality does not depend on $m$ and $t \in\left[0, T_{m}\right)$, and therefore

$$
\sup _{m \in \mathbb{N}} \sup _{t \in\left[0, T_{m}\right)}\left\|u_{m}(t)\right\|_{H}^{2}<\infty
$$

Inserting this estimate into (8.9), we obtain the existence of a constant $C_{5} \geq 0$ which is independent of $m$ and $t$ such that

$$
\begin{equation*}
\frac{c_{1}}{2} \int_{0}^{t}\left\|\dot{u}_{m}(s)\right\|_{H}^{2} d s+\left\|u_{m}(t)\right\|_{H}^{2}+\mathscr{E}_{\omega}\left(u_{m}(t)\right) \leq C_{5} \tag{8.10}
\end{equation*}
$$

Since the first two terms on the left-hand side are positive, this inequality implies that the set $\left\{u_{m}(t): m \in \mathbb{N}, t \in\left(0, T_{m}\right)\right\}$ is contained in the sublevel set $K_{C_{5}}$ of $\mathscr{E}_{\omega}$. By coercivity of $\mathscr{E}_{\omega}, K_{C_{5}}$ is bounded in $V$, and therefore

$$
\sup _{m \in \mathbb{N}} \sup _{t \in\left[0, T_{m}\right)}\left\|u_{m}(t)\right\|_{V}<\infty .
$$

We use again the fact that $\mathscr{E}_{\omega}$ is bounded from below and deduce in addition from inequality (8.10) that

$$
\sup _{m \in \mathbb{N}}\left\|u_{m}\right\|_{W^{1,2}\left(0, T_{m} ; H\right)}<\infty
$$

Since $T_{m} \leq T$ is finite, this implies that for each $m \in \mathbb{N}$ the function $\dot{u}_{m}$ is integrable on $\left[0, T_{m}\right)$. Hence, $u_{m}$ extends to a continuous function on the closed interval $\left[0, T_{m}\right]$, and Carathéodory's theorem and the definition of maximal solution imply that this is only possible if $T_{m}=T$, that is, the solutions $u_{m}$ are global.

From the preceding two inequalities and the continuous embedding $V \hookrightarrow H$ we obtain that

$$
\begin{equation*}
\left(u_{m}\right) \text { is bounded in } W^{1,2}(0, T ; H) \cap L^{\infty}(0, T ; V) . \tag{8.11}
\end{equation*}
$$

By assumption, the derivative $\mathscr{E}^{\prime}$ maps bounded sets into bounded sets, so that the boundedness of $\left(u_{m}\right)$ in $L^{\infty}(0, T ; V)$ implies that

$$
\left(\mathscr{E}^{\prime}\left(u_{m}\right)\right) \text { is bounded in } L^{\infty}\left(0, T ; V^{\prime}\right) .
$$

Part 3 (Extracting a convergent subsequence): The space $W^{1,2}(0, T ; H)$ is a Hilbert space and the spaces $L^{\infty}(0, T ; V) \cong L^{1}\left(0, T ; V^{\prime}\right)^{\prime}$ and $L^{\infty}\left(0, T ; V^{\prime}\right) \cong$ $L^{1}(0, T ; V)^{\prime}$ are dual space by reflexivity of $V$ (and $V^{\prime}$ ) and by Theorem 5.6. Moreover, the spaces $L^{1}\left(0, T ; V^{\prime}\right)$ and $L^{1}(0, T ; V)$ are separable by Theorem 5.5. Hence, by Theorem 6.3, by the assumption that $V$ is compactly embedded into $H$, and by Lemma 8.5 , there exist $u \in W^{1,2}(0, T ; H), v \in L^{\infty}(0, T ; V), w \in C([0, T] ; H)$, $\chi \in L^{\infty}\left(0, T ; V^{\prime}\right)$ and a subsequence of $\left(u_{m}\right)$ (which we denote for simplicity again by $\left(u_{m}\right)$ ) such that

$$
\begin{align*}
& u_{m} \rightharpoonup u \text { in } W^{1,2}(0, T ; H), \\
& u_{m} \xrightarrow{\text { weak* }} v \text { in } L^{\infty}(0, T ; V), \\
& u_{m} \rightarrow w \text { in } C([0, T] ; H), \text { and }  \tag{8.12}\\
& \mathscr{E}^{\prime}\left(u_{m}\right) \xrightarrow{\text { weak* }} \chi \text { in } L^{\infty}\left(0, T ; V^{\prime}\right) .
\end{align*}
$$

We leave it as an exercise to show that $u \in W^{1,2}(0, T ; H) \cap L^{\infty}(0, T ; V)$ and $u=v=w$. Moreover, it is an exercise to show that every continuous linear operator between two Hilbert spaces maps weakly convergent sequences into weakly convergent sequences. Since the operator $W^{1,2}(0, T ; H) \rightarrow L^{2}(0, T ; H), u \mapsto \dot{u}$ is continuous and linear, and by (8.12),

$$
\begin{aligned}
& \dot{u}_{m} \rightharpoonup \dot{u} \text { in } L^{2}(0, T ; H), \\
& u_{m}(0) \rightarrow u(0) \text { in } H \text { and } \\
& u_{m}(T) \rightarrow u(T) \text { in } H .
\end{aligned}
$$

The above weak respectively weak* convergences in the respective function spaces mean that

$$
\begin{align*}
& \int_{0}^{T}\left\langle v, u_{m}\right\rangle_{V^{\prime}, V} \rightarrow \int_{0}^{T}\langle v, u\rangle_{V^{\prime}, V} \text { for every } v \in L^{1}\left(0, T ; V^{\prime}\right), \\
& \int_{0}^{T}\left\langle\dot{u}_{m}, v\right\rangle_{H} \rightarrow \int_{0}^{T}\langle\dot{u}, v\rangle_{H} \text { for every } v \in L^{2}(0, T ; H) \text {, and }  \tag{8.13}\\
& \int_{0}^{T}\left\langle\mathscr{E}^{\prime}\left(u_{m}\right), v\right\rangle_{V^{\prime}, V} \rightarrow \int_{0}^{T}\langle\chi, v\rangle_{V^{\prime}, V} \text { for every } v \in L^{1}(0, T ; V) .
\end{align*}
$$

Part 4 (Showing that the limit $u$ is a solution): First of all, we have just seen that $u_{m}(0) \rightarrow u(0)$ in $H$. On the other hand, $u_{m}(0)=u_{0}^{m}$ by (8.6), and $u_{0}^{m} \rightarrow u_{0}$ in $V$ by the choice of the sequence $\left(u_{0}^{m}\right)$. Since $V$ is continuously embedded into $H$, we obtain $u(0)=u_{0}$, that is, $u$ satisfies the initial condition in (8.4). It remains to show that $u$ satisfies also the differential equation.

Let $w \in V_{m}$ and $\varphi \in L^{2}(0, T)$. For every $n \geq m$ we multiply equation (8.7) by $\varphi(\cdot) w$ with respect to the inner product $\langle\cdot \cdot \cdot\rangle_{g\left(u_{n}\right)}$, integrate over $(0, T)$ and obtain that

$$
\begin{aligned}
& \int_{0}^{T}\left\langle\dot{u}_{n}(t), \varphi(t) w\right\rangle_{g\left(u_{n}(t)\right)} d t+\int_{0}^{T}\left\langle\mathscr{E}^{\prime}\left(u_{n}(t)\right), \varphi(t) w\right\rangle_{V^{\prime}, V} d t= \\
&=\int_{0}^{T}\langle f(t), \varphi(t) w\rangle_{g\left(u_{n}(t)\right)} d t
\end{aligned}
$$

Letting $n \rightarrow \infty$ in this last equality and using (8.13) and the continuity assumption on $g$, we obtain

$$
\int_{0}^{T}\langle\dot{u}(t), \varphi(t) w\rangle_{g(u(t))} d t+\int_{0}^{T}\langle\chi(t), \varphi(t) w\rangle_{V^{\prime}, V} d t=\int_{0}^{T}\langle f(t), \varphi(t) w\rangle_{g(u(t))} d t
$$

Using the fact that $\left\{\varphi(\cdot) w: w \in \bigcup_{m} V_{m}, \varphi \in L^{2}(0, T)\right\}$ spans a dense subspace of $L^{2}(0, T ; V)$ (Theorem 5.5), we obtain for every $v \in L^{2}(0, T ; V)$

$$
\begin{equation*}
\int_{0}^{T}\langle\dot{u}, v\rangle_{g(u)}+\int_{0}^{T}\langle\chi, v\rangle_{V^{\prime}, V}=\int_{0}^{T}\langle f, v\rangle_{g(u)} . \tag{8.14}
\end{equation*}
$$

It is left to show that $\chi=\mathscr{E}^{\prime}(u)$. We multiply equation (8.7) by $u_{m}$ with respect to the inner product $\langle\cdot, \cdot\rangle_{g\left(u_{m}\right)}$, integrate the result over $(0, T)$, and obtain

$$
\begin{equation*}
\int_{0}^{T}\left\langle\mathscr{E}^{\prime}\left(u_{n}\right), u_{n}\right\rangle_{V^{\prime}, V}=\int_{0}^{T}\left\langle f, u_{n}\right\rangle_{g\left(u_{n}\right)}-\int_{0}^{T}\left\langle\dot{u}_{n}, u_{n}\right\rangle_{g\left(u_{n}\right)} . \tag{8.15}
\end{equation*}
$$

The continuity assumption (8.3) on $g$ implies

$$
\int_{0}^{T}\left\langle f, u_{n}\right\rangle_{g\left(u_{n}\right)} \longrightarrow \int_{0}^{T}\langle f, u\rangle_{g(u)} .
$$

Moreover, the continuity assumption on $g$, the uniform convergence of ( $u_{n}$ ) (see (8.12)), the Cauchy-Schwarz inequality and the uniform equivalence of $\langle\cdot, \cdot\rangle_{g(u)}$ and
$\langle\cdot, \cdot\rangle_{H}$ imply

$$
\begin{aligned}
\int_{0}^{T}\left\langle\dot{u}_{n}, u_{n}\right\rangle_{g\left(u_{n}\right)} & =\int_{0}^{T}\left\langle\dot{u}_{n}, u\right\rangle_{g\left(u_{n}\right)}+\int_{0}^{T}\left\langle\dot{u}_{n}, u_{n}-u\right\rangle_{g\left(u_{n}\right)} \\
& \rightarrow \int_{0}^{T}\langle\dot{u}, u\rangle_{g(u)} .
\end{aligned}
$$

Hence, if we let $n \rightarrow \infty$ in equation (8.15) and if we use the equality (8.14) with $v=u$, we obtain

$$
\int_{0}^{T}\left\langle\mathscr{E}^{\prime}\left(u_{n}\right), u_{n}\right\rangle_{V^{\prime}, V} d t \longrightarrow \int_{0}^{T}\langle\chi, u\rangle_{V^{\prime}, V} d t
$$

Since $\mathscr{E}_{\omega}^{\prime}(u)=\mathscr{E}^{\prime}(u)+\omega u$, since $u_{n} \rightarrow u$ in $C([0, T] ; H)$ (see (8.12)), and since $\left\langle u_{n}, u_{n}\right\rangle_{V^{\prime}, V}=\left\|u_{n}\right\|_{H}^{2}$, this last convergence implies

$$
\begin{equation*}
\int_{0}^{T}\left\langle\mathscr{E}_{\omega}^{\prime}\left(u_{n}\right), u_{n}\right\rangle_{V^{\prime}, V} \longrightarrow \int_{0}^{T}\langle\chi+\omega u, u\rangle_{V^{\prime}, V} \tag{8.16}
\end{equation*}
$$

Let $v \in L^{\infty}(0, T ; V)$ and $\lambda \in \mathbb{R}$. By Lemma 6.2 (applied to the convex function $\mathscr{E}_{\omega}$ ) and integration over $(0, T)$,

$$
\int_{0}^{T}\left\langle\mathscr{E}_{\boldsymbol{\omega}}^{\prime}\left(u_{m}\right), u_{m}-u-\lambda v\right\rangle_{V^{\prime}, V} \geq \int_{0}^{T}\left\langle\mathscr{E}_{\boldsymbol{\omega}}^{\prime}(u+\lambda v), u_{m}-u-\lambda v\right\rangle_{V^{\prime}, V}
$$

Letting $m \rightarrow \infty$ in this inequality, we obtain on using again the weak* convergences $u_{m} \xrightarrow{\text { weak* }} u, \mathscr{E}_{\omega}^{\prime}\left(u_{m}\right) \xrightarrow{\text { weak* }} \chi+\omega u$, and (8.16) that

$$
-\int_{0}^{T}\langle\chi+\omega u, \lambda v\rangle_{V^{\prime}, V} \geq-\int_{0}^{T}\left\langle\mathscr{E}_{\omega}^{\prime}(u+\lambda v), \lambda v\right\rangle_{V^{\prime}, V}
$$

We divide by $\lambda>0$ and by $\lambda<0$, let $\lambda \rightarrow 0+$ and $\lambda \rightarrow 0-$, and use the continuity of $\mathscr{E}$ in order to obtain that

$$
\int_{0}^{T}\langle\chi, v\rangle_{V^{\prime}, V}=\int_{0}^{T}\left\langle\mathscr{E}^{\prime}(u), v\right\rangle_{V^{\prime}, V}
$$

Since $v \in L^{\infty}(0, T ; V)$ is arbitrary, this implies

$$
\mathscr{E}^{\prime}(u)=\chi
$$

Hence, we may replace $\chi$ by $\mathscr{E}^{\prime}(u)$ in equality (8.14) and we deduce that $u$ is a solution of the variational form of $\dot{u}+\nabla_{g} \mathscr{E}(u)=f$ on $[0, T]$. In particular, $u$ is a solution of the gradient system (8.4).

## Uniqueness and energy inequality

See Exercise 8.1 (c).

### 8.2 Exercises

8.1. a) Prove that a constant metric $g: V \rightarrow \operatorname{Inner}(H)$ - that is, $\langle\cdot, \cdot\rangle_{g(u)}=\langle\cdot, \cdot\rangle_{H}$ for every $u \in V$ - satisfies the continuity condition (8.3) from Theorem 8.1.
b) Show that if a function $g: V \rightarrow \operatorname{Inner}(H)$ satisfies the continuity condition (8.3), then it is continuous in the following sense:

$$
\left.\begin{array}{l}
u_{n} \rightharpoonup u \text { in } V,  \tag{8.17}\\
v_{n} \rightharpoonup v \text { in } H, \text { and } \\
w \in H
\end{array}\right\} \Rightarrow\left\langle v_{n}, w\right\rangle_{g\left(u_{n}\right)} \rightarrow\langle v, w\rangle_{g(u)}
$$

In particular, $g$ is a metric, that is, $u_{n} \rightarrow u$ in $V$ implies $\langle v, w\rangle_{g\left(u_{n}\right)} \rightarrow\langle v, w\rangle_{g(u)}$ for every $v, w \in H$.
Open problem. Prove or disprove that the condition (8.17) implies (8.3).
c) Prove the addendum in Theorem 8.1, that is, prove uniqueness of solutions and the energy inequality (8.5).
8.2 (Quadratic forms). Let $\mathscr{E}: V \rightarrow \mathbb{R}$ be a continuous quadratic form on a Banach space $V$ which is densely and continuously embedded into a Hilbert space $H$.
a) Show that if $\mathscr{E}$ is coercive, then $\mathscr{E}$ is convex.
b) Show that $\mathscr{E}$ is $H$-elliptic if and only if there exist $\omega \geq 0, \eta>0$ such that

$$
\mathscr{E}(u)+\frac{\omega}{2}\|u\|_{H}^{2} \geq \eta\|u\|_{V}^{2} \quad \text { for every } u \in V
$$

c) Show that if $\mathscr{E}$ is $H$-elliptic, then there exists an equivalent norm on $V$ which comes from an inner product (that is, $V$ is a Hilbert space).
d) Show that if $\mathscr{E}$ is $H$-elliptic, then

$$
\|u\|_{D\left(\nabla_{H} \mathscr{E}\right)}^{2}:=\|u\|_{H}^{2}+\left\|\nabla_{H} \mathscr{E}(u)\right\|_{H}^{2}
$$

defines a complete norm on the domain $D\left(\nabla_{H} \mathscr{E}\right)$.
e) Show that if $\mathscr{E}$ is $H$-elliptic, then for every $u_{0} \in V$ and every $f \in L^{2}(0, T ; H)$ the gradient system

$$
\dot{u}+\nabla_{H} \mathscr{E}(u)=f, \quad u(0)=u_{0}
$$

admits a unique solution $u \in H^{1}(0, T ; H) \cap L^{\infty}(0, T ; V) \cap L^{2}\left(0, T ; D\left(\nabla_{H} \mathscr{E}\right)\right)$. Hint. We do not assume that the embedding $V \hookrightarrow H$ is compact so that The-
orem 8.1 does not apply. Use Theorem 6.1 and the observation that $u$ is a solution if and only if $v(t)=e^{-\omega t} u(t)$ is a solution of

$$
\dot{v}+\nabla_{H} \mathscr{E}(v)+\omega v=e^{-\omega t} f, \quad v(0)=u_{0} .
$$

8.3 (Perona-Malik model). Let $\Omega \subseteq \mathbb{R}^{d}$ be open and bounded. Let $c: \mathbb{R}_{+} \rightarrow(0, \infty)$ be a continuously differentiable function such that there exist $p>1, c_{1} \geq 0$ and $s_{0} \geq 0$ such that $c(s) s \leq c_{1} s^{p-1}$ for every $s \geq s_{0}$. Given $f \in L^{2}(\Omega)$, we call a function $u \in W^{1, p}(\Omega) \cap L^{2}(\Omega)$ a weak solution of the problem

$$
\begin{cases}-\operatorname{div}(c(|\nabla u|) \nabla u)=f & \text { in } \Omega  \tag{8.18}\\ \frac{\partial u}{\partial \mathbf{n}}=0 & \text { on } \partial \Omega\end{cases}
$$

if for every $\varphi \in W^{1, p}(\Omega) \cap L^{2}(\Omega)$ one has

$$
\int_{\Omega} c(|\nabla u|) \nabla u \nabla \varphi=\int_{\Omega} f \varphi .
$$

a) Assume in addition that $\Omega$ is of class $C^{1}$ and $f \in C(\bar{\Omega})$. Show that $u \in C^{2}(\bar{\Omega})$ is a weak solution of the problem (8.18) if and only if $u$ is a classical solution, that is, all derivatives are understood to be classical derivatives and $u$ satisfies the partial differential equation and the boundary condition in the usual sense.
b) Let $c(s)=(1+s)^{p-2}$ with $p \in(1,2)$, so that $c$ satisfies the properties from the Perona-Malik model. Let $C(s):=\int_{0}^{s} c(r) r d r$. Consider the Banach space $V=W^{1, p}(\Omega) \cap L^{2}(\Omega)$ with norm $\|u\|_{V}=\|u\|_{W^{1, p}}+\|u\|_{L^{2}}$. Show that the function $\mathscr{E}: V \rightarrow \mathbb{R}, \mathscr{E}(u)=\int_{\Omega} C(|\nabla u|)$ is continuously differentiable, $L^{2}(\Omega)$ elliptic, and that $\mathscr{E}$ ' maps bounded sets into bounded sets. Show that $u$ is a weak solution of (8.18) if and only if $\nabla_{L^{2}} \mathscr{E}(u)=f$.
Remark. Recall from Exercise 7.2 that $\mathscr{E}$ is convex if $C$ is convex and nondecreasing. Note that $C$ is convex if and only if $s c(s)$ is nondecreasing. In particular, if $C$ is convex, then the coefficient $c$ can not decrease arbitrarily rapidly.
c) In addition to the assumptions of (b), assume that $\Omega \subseteq \mathbb{R}^{2}$ is of class $C^{1}$. Conclude that for every $u_{0} \in W^{1, p}(\Omega)$ and every $f \in L^{2}\left(0, T ; L^{2}(\Omega)\right)$ there exists a unique solution $u \in W^{1,2}\left(0, T ; L^{2}(\Omega)\right) \cap L^{\infty}\left(0, T ; W^{1, p}(\Omega)\right)$ of the problem

$$
\dot{u}+\nabla_{L^{2}} \mathscr{E}(u)=f, \quad u(0)=u_{0}
$$

Write the variational form of this gradient system.
Remark. If $p \in(1,2)$ and $\Omega \subseteq \mathbb{R}^{2}$ is bounded and of class $C^{1}$, then, by the Rellich-Kondrachev theorem, $W^{1, p}(\Omega) \hookrightarrow L^{2}(\Omega)$ with compact embedding.

## Lecture 9 <br> Regularity of solutions

After having proved existence of solutions of abstract gradient systems, several questions about their qualitative behaviour arise naturally. Some of them are about regularity of the solutions. Is the regularity of the solutions which we obtained in Theorems 6.1 and 8.1 the optimal one or can we expect higher regularity if the data in the gradient system (6.3) and (8.4) are more regular? For example, when can we expect that solutions of gradient systems are continuous or continuously differentiable with values in the energy space $V$ ? When does the energy inequality (8.5) hold, and when is it actually an energy equality? Finding answers to these questions is a crucial step before studying, for example, the long time behaviour of solutions of gradient systems.

It turns out that for a certain class of gradient systems, including linear, and so-called semilinear and quasilinear problems, higher regularity can be obtained from mere existence and maximal regularity results such as Theorems 6.1 and 8.1. The idea of the proof, which we present in this lecture, is the so-called parameter trick and a clever application of the implicit function theorem; it is due to Angenent (see [Angenent (1990b)] and also [Angenent (1990a)]).

The regularity result in this lecture is true for abstract differential equations and there is no necessity to confine the study to gradient systems. We come back to gradient systems in the following lectures.

### 9.1 Interpolation

In this section, we recall shortly some results which were the subject of Exercise 5.2. Let $D$ and $X$ be two Banach spaces such that $D$ is densely and continuously embedded into $X$. For every $p \in[1, \infty)$ we consider the space

$$
\operatorname{MR}_{p}(X, D):=W^{1, p}(0,1 ; X) \cap L^{p}(0,1 ; D)
$$

equipped with the norm

$$
\|u\|_{\mathrm{MR}_{p}(X, D)}^{p}:=\|u\|_{W^{1, p}(0, T ; X)}^{p}+\|u\|_{L^{p}(0, T ; D)}^{p},
$$

which turns it into a Banach space. We recall from Theorem 5.10 that every function $u \in \operatorname{MR}_{p}(X, D)$ is continuous with values in $X$; more precisely, every element $u \in \operatorname{MR}_{p}(X, D)$ admits a representative which is continuous with values in $X$. It makes therefore sense to evaluate functions $u \in \operatorname{MR}_{p}(X, D)$ in every $t \in[0,1]$ and in particular to define the trace space

$$
\operatorname{Tr}_{p}(X, D):=\left\{x \in X: \text { there exists } u \in \operatorname{MR}_{p}(X, D) \text { such that } u(0)=x\right\}
$$

This space is a Banach space for the norm

$$
\|x\|_{\operatorname{Tr}_{p}}:=\inf \left\{\|u\|_{\mathrm{MR}_{p}}: u \in \operatorname{MR}_{p}(X, D) \text { and } u(0)=x\right\}
$$

The fact that $\operatorname{Tr}_{p}(X, D)$ is a Banach space becomes clear from considering the commutative diagram

in which $\delta_{0}$ is the point evaluation at $t=0$ (a continuous linear operator, by Theorem 5.10), $\operatorname{MR}_{p}^{0}(X, D):=\left\{u \in \operatorname{MR}_{p}(X, D): u(0)=0\right\}$ is its kernel (a closed subspace, by continuity of $\delta_{0}$ ), $q$ is the canonical quotient map, $b$ is the canonical bijection onto the range of $\delta_{0}$, and $i$ is the canonical injection. The quotient space $\operatorname{MR}_{p}(X, D) / \operatorname{MR}_{p}^{0}(X, D)$ is clearly a Banach space and $\|\cdot\|_{\operatorname{Tr}_{p}}$ is the natural norm which turns $b$ into an isometric isomorphism. In the literature, the trace space $\operatorname{Tr}_{p}(X, D)$ is also denoted by $(X, D)_{\frac{1}{p^{\prime}}, p}$ (with $p^{\prime}=\frac{p}{p-1}$ ); it is a particular example in the scale of so-called real interpolation spaces $(X, D)_{\theta, p}$ which depend on two parameters $(0<\theta<1,1 \leq p \leq \infty$ or $(\theta, p)=(1, \infty)$ ); see [Lunardi (1995), Chapter 1], [Lunardi (2009)]. One has

$$
D \subseteq \operatorname{Tr}_{p}(X, D) \subseteq X
$$

and the embeddings are dense and continuous. It is also true - but a little bit less obvious to prove - that

$$
\operatorname{MR}_{p}(X, D) \subseteq C\left([0,1] ; \operatorname{Tr}_{p}(X, D)\right)
$$

with dense and continuous embedding. Using translations and dilations, it is clear from this embedding that for every $a, b \in \mathbb{R}, a<b$, one has

$$
\begin{equation*}
W^{1, p}(a, b ; X) \cap L^{p}(a, b ; D) \hookrightarrow C\left([a, b] ; \operatorname{Tr}_{p}(X, D)\right) . \tag{9.1}
\end{equation*}
$$

### 9.2 Time regularity - Angenent's trick

Let $D$ and $X$ be two Banach spaces such that $D$ is densely and continuously embedded into $X$. Let $F: D \rightarrow X$ be a continuously differentiable function and let $f:[0, T] \rightarrow X$ be an integrable function. In this section we study the regularity of solutions of the abstract differential equation

$$
\begin{equation*}
\dot{u}+F(u)=f \tag{9.2}
\end{equation*}
$$

Theorem 9.1 (Time regularity). Let $p \in[1, \infty)$ and $0<T<T^{\prime}$. Assume that, for some $k \geq 1$, the composition operator $\mathscr{F}: W^{1, p}(0, T ; X) \cap L^{p}(0, T ; D) \rightarrow$ $L^{p}(0, T ; X), u \mapsto F(u)=F \circ u$ is $k$ times continuously differentiable. Let $f \in$ $W^{k, p}\left(0, T^{\prime} ; X\right)$, and let $u \in W^{1, p}\left(0, T^{\prime} ; X\right) \cap L^{p}\left(0, T^{\prime} ; D\right)$ be a solution of $(9.2)$ in the sense that the equality holds almost everywhere on $\left(0, T^{\prime}\right)$. Assume that the linear problem

$$
\left\{\begin{array}{l}
\dot{v}+F^{\prime}(u) v=g  \tag{9.3}\\
v(0)=0
\end{array}\right.
$$

admits for every $g \in L^{p}(0, T ; X)$ a unique solution $v \in W^{1, p}(0, T ; X) \cap L^{p}(0, T ; D)$. Then

$$
\begin{aligned}
& u \in W_{l o c}^{k+1, p}((0, T] ; X) \cap W_{l o c}^{k, p}((0, T] ; D) \cap C^{k}\left((0, T] ; \operatorname{Tr}_{p}(X, D)\right) \text { and } \\
& t \mapsto t^{j} u^{(j)}(t) \in W^{1, p}(0, T ; X) \cap L^{p}(0, T ; D) \quad \text { for every } j=0, \ldots, k
\end{aligned}
$$

If $\mathscr{F}$ and $f$ are of class $C^{\infty}$, then $u \in C^{\infty}((0, T] ; D)$.
Observe that we assume no higher regularity of the function $F$, but we rather assume higher regularity of the composition operator $\mathscr{F}$. Of course, one then presumes higher regularity of the function $F$, but note carefully that $F \in C^{k}$ alone does in general not imply $\mathscr{F} \in C^{k}$. The implication " $F \in C^{k} \Rightarrow \mathscr{F} \in C^{k}$ " is true in special situations, for example if one assumes additional growth conditions on the derivatives of $F$; see also Theorem 4.3. Here are two natural cases in which the regularity assumption on $\mathscr{F}$ is satisfied (see Exercises).
a) The function $F: D \rightarrow X$ is continuous and linear.
b) The function $F: D \rightarrow X$ extends to a $k$ times continuously differentiable function from $\operatorname{Tr}_{p}(X, D)$ into $X$.
c) The function $F: D \rightarrow X$ is the sum of two functions as in (a) or (b).

Maximal regularity plays an important role in Theorem 9.1, in fact in two ways. On the one hand, there is the maximal regularity assumption on the linear, nonautonomous problem (9.3) (the linear operator $F^{\prime}(u)=F^{\prime} \circ u$ depends in fact on time): for every given right-hand side $g \in L^{p}(0, T ; X)$ of (9.3) there exists a unique solution $v$ such that the two terms on the left-hand side have the same regularity as $g$. On the other hand, the existence of the solution $u$ of the problem (9.2), which is assumed in Theorem 9.1, is in concrete applications obtained by some existence and
maximal regularity result for which only $f \in L^{p}(0, T ; X)$ was assumed: two such results are our Theorems 6.1 and 8.1 which are existence and maximal regularity results for gradient systems. The maximal regularity assumptions in Theorem 9.1 are the optimal assumptions from the point of view of the proof in which we apply the following classical theorem from calculus (see Theorem C. 6 for the proof).

Theorem 9.2 (Implicit function theorem). Let $X_{1}, X_{2}$ and $Y$ be three Banach spaces, and let $U \subseteq X_{1} \times X_{2}$ be an open set. Let $G: U \rightarrow Y$ be continuously differentiable. Let $\bar{x}=\left(\bar{x}_{1}, \bar{x}_{2}\right) \in U$ be such that $\frac{\partial G}{\partial x_{2}}(\bar{x}): X_{2} \rightarrow Y$ is an isomorphism. Then there exist neighbourhoods $U_{1} \subseteq X_{1}$ of $\bar{x}_{1}$ and $U_{2} \subseteq X_{2}$ of $\bar{x}_{2}, U_{1} \times U_{2} \subseteq U$, and a continuously differentiable function $g: U_{1} \rightarrow U_{2}$ such that

$$
\left\{\left(x_{1}, x_{2}\right) \in U_{1} \times U_{2}: G\left(x_{1}, x_{2}\right)=G\left(\bar{x}_{1}, \bar{x}_{2}\right)\right\}=\left\{\left(x_{1}, g\left(x_{1}\right)\right): x_{1} \in U_{1}\right\}
$$

If, in addition, $G$ is $k$ times continuously differentiable, then the implicit function $g$ is $k$ times continuously differentiable, too.

Lemma 9.3. Let $A: W^{1, p}(0, T ; X) \cap L^{p}(0, T ; D) \rightarrow L^{p}(0, T ; X)$ be a continuous, linear operator. Assume that the problem

$$
\left\{\begin{array}{l}
\dot{v}+A v=g  \tag{9.4}\\
v(0)=0,
\end{array}\right.
$$

admits for every $g \in L^{p}(0, T ; X)$ a unique solution $v \in W^{1, p}(0, T ; X) \cap L^{p}(0, T ; D)$. Then, for every $g \in L^{p}(0, T ; X)$ and every $v_{0} \in \operatorname{Tr}_{p}(X, D)$ the problem

$$
\left\{\begin{array}{l}
\dot{v}+A v=g  \tag{9.5}\\
v(0)=v_{0},
\end{array}\right.
$$

admits a unique solution $v \in W^{1, p}(0, T ; X) \cap L^{p}(0, T ; D)$.
Proof. Uniqueness of solutions of (9.5) follows from linearity and uniqueness of solutions of (9.4). In order to show existence, let $g \in L^{p}(0, T ; X)$ and $v_{0} \in \operatorname{Tr}_{p}(X, D)$. By definition of the trace space (we use in fact the definition and a dilation of the interval $(0,1)$ onto $(0, T)$ ), there exists $w \in W^{1, p}(0, T ; X) \cap L^{p}(0, T ; D)$ such that $w(0)=v_{0}$. For this function $w$ one has $\dot{w}+A w \in L^{p}(0, T ; X)$. By assumption, there exists a unique $z \in W^{1, p}(0, T ; X) \cap L^{p}(0, T ; D)$ which is solution of

$$
\left\{\begin{array}{l}
\dot{z}+A z=g-\dot{w}-A w, \\
z(0)=0 .
\end{array}\right.
$$

Setting $v=w+z$, we obtain a solution of (9.5).
Proof (Proof of Theorem 9.1). It is useful to set

$$
\operatorname{MR}_{p}(0, T ; X, D):=W^{1, p}(0, T ; X) \cap L^{p}(0, T ; D)
$$

Let $\varepsilon \in\left(0, \frac{T^{\prime}-T}{T}\right)$. For every $\lambda \in(-\varepsilon, \varepsilon)$ and every $t \in(0, T)$ we define

$$
u_{\lambda}(t):=u(t+\lambda t) .
$$

Then, for every $\lambda \in(-\varepsilon, \varepsilon)$ the function $u_{\lambda}$ belongs to $\operatorname{MR}_{p}(0, T ; X, D)$ and it solves the nonlinear problem

$$
\left\{\begin{array}{l}
\dot{u}_{\lambda}+(1+\lambda) F\left(u_{\lambda}\right)=(1+\lambda) f((1+\lambda) \cdot),  \tag{9.6}\\
u_{\lambda}(0)=u(0) .
\end{array}\right.
$$

Starting from this observation, we consider the nonlinear operator

$$
\begin{aligned}
& G:(-\varepsilon, \varepsilon) \times \operatorname{MR}_{p}(0, T ; X, D) \rightarrow L^{p}(0, T ; X) \times \operatorname{Tr}_{p}(X, D), \\
& \quad(\lambda, v) \mapsto(\dot{v}+(1+\lambda) F(v)-(1+\lambda) f((1+\lambda) \cdot), v(0)-u(0)) .
\end{aligned}
$$

It follows from the assumptions $-\mathscr{F}$ is $k$ times continuously differentiable and $f \in W^{k, p}\left(0, T^{\prime} ; X\right)$-, that the operator $G$ is $k$ times continuously differentiable. Moreover, by definition of the operator $G$, and since the functions $u_{\lambda}$ are solutions of (9.6), one has

$$
G\left(\lambda, u_{\lambda}\right)=(0,0) \text { for every } \lambda \in(-\varepsilon, \varepsilon)
$$

We show that $G$ satisfies the assumptions of the implicit function theorem at $(0, u)=\left(0, u_{0}\right)$. For this, we have to consider the partial derivative $\frac{\partial G}{\partial v}$ at $(0, u)$ which is the linear operator given by

$$
\begin{aligned}
\frac{\partial G}{\partial v}(0, u): \operatorname{MR}_{p}(0, T ; X, D) & \rightarrow L^{p}(0, T ; X) \times \operatorname{Tr}_{p}(X, D), \\
v & \mapsto\left(\dot{v}+F^{\prime}(u) v, v(0)\right)
\end{aligned}
$$

Observe at this point that, by our assumption on the linear problem (9.3) and by Lemma 9.3, the problem

$$
\left\{\begin{array}{l}
\dot{v}+F^{\prime}(u) v=g \\
v(0)=v_{0}
\end{array}\right.
$$

admits for every $g \in L^{p}(0, T ; X)$ and every $v_{0} \in \operatorname{Tr}_{p}(X, D)$ a unique solution $v \in \operatorname{MR}_{p}(0, T ; X, D)$. In other words, the operator $\frac{\partial G}{\partial v}(0, u)$ is bijective. By the bounded inverse theorem, the linear operator $\frac{\partial G}{\partial v}(0, u)$ is continuously invertible. Hence, by the implicit function theorem, there exists $\varepsilon^{\prime} \in(0, \varepsilon)$, a neighbourhood $U \subseteq \operatorname{MR}_{p}(0, T ; X, D)$ of $u$, and a $k$ times continuously differentiable implicit function $g:\left(-\varepsilon^{\prime}, \varepsilon^{\prime}\right) \rightarrow U$ such that

$$
G(\lambda, g(\lambda))=G\left(0, u_{0}\right)=(0,0)
$$

Moreover, all solutions in $\left(-\varepsilon^{\prime}, \varepsilon^{\prime}\right) \times U$ of the equation $G(\lambda, v)=(0,0)$ are of the form $(\lambda, g(\lambda))$. Since, for every $\lambda \in\left(-\varepsilon^{\prime}, \varepsilon^{\prime}\right)$, the elements $\left(\lambda, u_{\lambda}\right)$ are solutions of this equation, we obtain that $u_{\lambda}=g(\lambda)$. In particular, we have obtained that the
function

$$
\begin{aligned}
g:\left(-\varepsilon^{\prime}, \varepsilon^{\prime}\right) & \rightarrow \operatorname{MR}_{p}(0, T ; X, D), \\
\lambda & \mapsto u_{\lambda}=u((1+\lambda) \cdot),
\end{aligned}
$$

is $k$ times continuously differentiable. This is the desired information. Since the point evaluation $\operatorname{MR}_{p}(0, T ; X, D) \rightarrow X, v \mapsto v(t)$ is linear, we obtain that the function $\lambda \rightarrow u(t+\lambda t)$ is $k$ times continuously differentiable with values in $X$. In particular, $u$ is $k$ times continuously differentiable with values in $X$ and $\left.\frac{d^{j}}{d \lambda^{j}} u(t+\lambda t)\right|_{\lambda=0}=$ $t^{j} u^{(j)}(t)$ for every $t \in(0, T)$ and every $j \in\{0, \ldots, k\}$. Coming back to the function $g$, we see that the consecutive derivatives of $g$ at $\lambda=0$ are given by

$$
\begin{aligned}
& t \mapsto t \dot{u}(t) \in \operatorname{MR}_{p}(0, T ; X, D) \\
& \vdots \\
& t \mapsto t^{k} u^{(k)}(t) \in \operatorname{MR}_{p}(0, T ; X, D)
\end{aligned}
$$

This is the stated regularity of the solution $u$.

### 9.3 Exercises

In order to formulate the exercises, consider first the following definition and the extension of the implicit function theorem (see [Zeidler (1990), Corollary 4.23]). Let $X$ and $Y$ be two Banach spaces, and let $U \subseteq X$ be an open set. We say that a function $F: U \rightarrow Y$ is analytic if it is infinitely many times differentiable, and if for every $x \in U$ there exists $r>0$ such that

$$
\begin{aligned}
& \sum_{k=0}^{\infty} \frac{1}{k!}\left\|F^{(k)}(x)\right\| r^{k}<\infty \text { and } \\
& F(x+h)=\sum_{k=0}^{\infty} \frac{1}{k!} F^{(k)}(x) h^{k} \quad \text { for every } h \in X \text { with }\|h\| \leq r \text { and } x+h \in U
\end{aligned}
$$

In this definition, $F^{(k)}(x)$ is the $k$-th derivative of $F$ at the point $x$. It is identified with a $k$-linear operator $X \times \cdots \times X \rightarrow Y$, and $F^{(k)}(x) h^{k}$ stands for $F^{(k)}(x)(h, \ldots, h)$ (the $k$-th derivative applied to the vector with $k$ identical entries $h$ ). For every $x \in U$ the supremum over all $r>0$ such that the above relations are true is called the radius of convergence of $F$ at $x$.

Theorem 9.4 (Addendum to the implicit function theorem, ). If, in the implicit function theorem 9.2 the function $G$ is analytic, then the implicit function $g$ is analytic, too.
9.1. a) By analyzing the proof of Theorem 9.1, by using the addendum to the Implicit function theorem, and by using the embedding (9.1), prove the following result.

Theorem 9.5 (Addendum to Theorem 9.1). If, in Theorem 9.1, the function $\mathscr{F}$ is analytic and $f=0$, then the solution $u:(0, T) \rightarrow \operatorname{Tr}_{p}(X, D)$ is analytic ( $u$ is even analytic with values in $D$ ). Morever, if $r(t)$ is the radius of convergence of $u$ at $t \in(0, T)$, then $r(t) \geq c t$ for some constant $c>0$ independent of $t$.
b) Every continuous linear operator is analytic. Starting from this observation, prove the following: Assume that $A: D \rightarrow X$ is a continuous, linear operator, and that the linear problem

$$
\left\{\begin{array}{l}
\dot{u}+A u=f  \tag{9.7}\\
u(0)=0
\end{array}\right.
$$

has $L^{p}$-maximal regularity in the sense that for every $f \in L^{p}(0, T ; X)$ it admits a unique solution $u \in W^{1, p}(0, T ; X) \cap L^{p}(0, T ; D)$. Show that then for every $u_{0} \in \operatorname{Tr}_{p}$ the initial value problem

$$
\left\{\begin{array}{l}
\dot{u}+A u=0,  \tag{9.8}\\
u(0)=u_{0}
\end{array}\right.
$$

admits a unique solution $u \in W^{1, p}(0, T ; X) \cap L^{p}(0, T ; D)$. This solution is analytic on $(0, T)$ with values in $\operatorname{Tr}_{p}$ and there exists $c>0$ such that the radius of convergence $r(t)$ of $u$ at $t$ satisfies $r(t) \geq c t$.
Remark for the reader who is familiar with the theory of linear $C_{0}$-semigroups. It has been shown in [Dore (1993)] that if the problem (9.7) has $L^{p}$-maximal regularity, then $-A$ with domain $D$ generates an analytic $C_{0}$-semigroup. Theorem 9.5 may be seen as a nonlinear variant of this result.
9.2 (Nemytski operators). Let $F: \mathbb{R} \rightarrow \mathbb{R}$ be continuously differentiable.
a) Show that for every compact $K \subseteq \mathbb{R}$ one has

$$
\lim _{h \rightarrow 0} \frac{F(s+h)-F(s)-F^{\prime}(s) h}{h}=0 \quad \text { uniformly in } s \in K .
$$

b) Let $\Omega \subseteq \mathbb{R}^{d}$ be an open set. Show that the composition operator $\mathscr{F}: L^{\infty}(\Omega) \rightarrow$ $L^{\infty}(\Omega), u \mapsto F(u)=F \circ u$ is continuously differentiable and that for every $u$, $h \in L^{\infty}(\Omega)$

$$
\mathscr{F}^{\prime}(u) h=F^{\prime}(u(\cdot)) h(\cdot) \quad \text { almost everywhere on } \Omega .
$$

Remark. The operator, which maps every function $m \in L^{\infty}(\Omega)$ to the multiplication operator $M \in \mathscr{L}\left(L^{\infty}(\Omega)\right)$ given by $M u(x):=m(x) u(x)\left(u \in L^{\infty}(\Omega)\right.$, $x \in \Omega)$, is isometric. We may therefore identify its range with $L^{\infty}(\Omega)$. In this
sense, we may identify the derivative $\mathscr{F}^{\prime}(u) \in \mathscr{L}\left(L^{\infty}(\Omega)\right)$ with the function $F^{\prime}(u) \in L^{\infty}(\Omega)$. With this identification, $\mathscr{F}^{\prime}$ is again a Nemytski operator.
c) Show that if $F$ is $k$ times continuously differentiable, then $\mathscr{F}$ is $k$ times continuously differentiable.
d) Show that if $F$ is analytic, then $\mathscr{F}$ is analytic, too.
9.3. Prove statements (a)-(c) on page 101 .
9.4. It may be instructive to consider the problem of regularity of solutions of the ordinary differential equation

$$
\dot{u}+F(u)=0
$$

where $F: \mathbb{R}^{d} \rightarrow \mathbb{R}^{d}$ is a continuous functions. Show by induction that if the function $F$ is $k$ times continuously differentiable ( $k \geq 0$ ), then every solution of this differential equation is $k+1$ times continuously differentiable (solution is meant here in the sense of Lectures 1 and 2. Compare your proof to the proof of Theorem 9.1. What about the case of analytic $F$ ?

## Lecture 10

## Existence of local solutions

In many textbooks on differential calculus, one can find the implicit function theorem and the local inverse theorem nearby. Depending on the particular textbook, one theorem is deduced from the other, and perhaps it is justified to call them brother and sister. Lecture 9 and the present lecture are somehow a copy of this reality in textbooks, because implicit function theorem and local inverse theorem appear nearby; we apply them both, in fact, to the same evolution equation.

Once again, in this lecture we turn our attention to the abstract differential equation

$$
\dot{u}+F(u)=f
$$

While in the preceding lecture we applied the implicit function theorem in order to prove regularity of solutions, in this lecture we prove existence of local solutions by the help of the local inverse theorem. It turns out that we meet in the assumptions of both theorems the same linear problem for which one has to prove existence, uniqueness and maximal regularity. Once one has existence, uniqueness and maximal regularity for the linear problem, existence of solutions is an immediate consequence of the local inverse theorem (Theorem 10.1 below).

In the second part of this lecture, we show how Theorems 9.1 and 10.1 may be applied to a semilinear diffusion equation, or, more precisely, to a gradient system associated with a semilinear diffusion equation. Due to the importance of an existence, uniqueness and maximal regularity result for a linear problem, Theorem 6.1 on the existence and uniqueness of global solutions of gradient systems becomes in this example an essential ingredient, although its application in the example of the semilinear diffusion equation is a little bit hidden. Find the two places where Theorem 6.1 is applied.

### 10.1 The local inverse theorem and existence of local solutions of abstract differential equations

Let $D$ and $X$ be two Banach spaces such that $D$ is densely and continuously embedded into $X$. Let $F: D \rightarrow X$ be a continuously differentiable function, $p \geq 1, T>0$, $f \in L^{p}(0, T ; X)$, and $u_{0} \in \operatorname{Tr}_{p}(X, D)$. We study existence of local solutions of the abstract differential equation

$$
\left\{\begin{array}{l}
\dot{u}+F(u)=f  \tag{10.1}\\
u(0)=u_{0}
\end{array}\right.
$$

As in the previous lecture, for simplicity of notation, we let

$$
\operatorname{MR}_{p}(0, T ; X, D):=W^{1, p}(0, T ; X) \cap L^{p}(0, T ; D)
$$

and we call a function $u \in \operatorname{MR}_{p}\left(0, T^{\prime} ; X, D\right)\left(0<T^{\prime} \leq T\right)$ a (local) solution if it satisfies the initial condition $u(0)=u_{0}$ and the differential equation $\dot{u}+F(u)=f$ almost everywhere on $\left(0, T^{\prime}\right)$.

Theorem 10.1 (Existence of local solutions). Assume that the composition operator $\operatorname{MR}_{p}(0, T ; X, D) \rightarrow L^{p}(0, T ; X), u \mapsto F(u)=F \circ u$ is everywhere defined and continuously differentiable. Assume further that there exists $\bar{u} \in \operatorname{MR}_{p}(0, T ; X, D)$ such that $\bar{u}(0)=u_{0}$ and for every $g \in L^{p}(0, T ; X)$ the linear problem

$$
\left\{\begin{array}{l}
\dot{v}+F^{\prime}(\bar{u}) v=g  \tag{10.2}\\
v(0)=0,
\end{array}\right.
$$

admits a unique solution $v \in \operatorname{MR}_{p}(0, T ; X, D)$. Then the problem (10.1) admits a local solution.

The proof of this theorem is almost an immediate consequence of the local inverse theorem (see Theorem C.5).

Theorem 10.2 (Local inverse theorem). Let $X$ and $Y$ be two Banach spaces and let $U \subseteq X$ be open. Let $G: U \rightarrow Y$ be a continuously differentiable function and let $\bar{x} \in U$ be such that $G^{\prime}(\bar{x}): X \rightarrow Y$ is an isomorphism, that is, continuous, linear, bijective and the inverse is also continuous. Then there exist neighbourhoods $V \subseteq U$ of $\bar{x}$ and $W \subseteq Y$ of $G(\bar{x})$ such that $G: V \rightarrow W$ is a $C^{1}$ diffeomorphism, that is $G$ is continuously differentiable, bijective and the inverse $G^{-1}: W \rightarrow V$ is continuously differentiable, too.

Proof (of Theorem 10.1). Consider the operator

$$
\begin{aligned}
G: \operatorname{MR}_{p}(0, T ; X, D) & \rightarrow L^{p}(0, T ; X) \times \operatorname{Tr}_{p}(X, D), \\
u & \mapsto(\dot{u}+F(u), u(0)) .
\end{aligned}
$$

By the assumption on $F$, and since the mappings $u \mapsto \dot{u}$ and $u \mapsto u(0)$ are continuous and linear, the operator $G$ is continuously differentiable. Its derivative at a point $u \in \operatorname{MR}_{p}(0, T ; X, D)$ is the linear operator given by

$$
\begin{aligned}
G^{\prime}(u): \operatorname{MR}_{p}(0, T ; X, D) & \rightarrow L^{p}(0, T ; X) \times \operatorname{Tr}_{p}(X, D), \\
v & \mapsto\left(\dot{v}+F^{\prime}(u) v, v(0)\right) .
\end{aligned}
$$

Observe that the assumption on the linear problem (10.2), together with Lemma 9.3, just translates the fact that $G^{\prime}(\bar{u})$ is bijective. Therefore, by the bounded inverse theorem, $G^{\prime}(\bar{u})$ is an isomorphism. By the local inverse theorem (Theorem 10.2), there exists a neighbourhood $V \subseteq \operatorname{MR}_{p}(0, T ; X, D)$ of $\bar{u}$ and a neighbourhood $W \subseteq L^{p}(0, T ; X) \times \operatorname{Tr}_{p}(X, D)$ of $G(\bar{u})=\left(\dot{\bar{u}}+F(\bar{u}), u_{0}\right)$ such that $G: V \rightarrow W$ is a $C^{1}$ diffeomorphism. In particular, for every $\left(h, z_{0}\right) \in W$ there exists a unique $z \in V$ such that

$$
\left\{\begin{array}{l}
\dot{z}+F(z)=h,  \tag{10.3}\\
z(0)=z_{0} .
\end{array}\right.
$$

For $T^{\prime} \in(0, T]$, consider the function $h \in L^{p}(0, T ; X)$ given by

$$
h(t)= \begin{cases}f(t) & \text { if } t \in\left(0, T^{\prime}\right) \\ \dot{\bar{u}}+F(\bar{u}) & \text { if } t \in\left(T^{\prime}, T\right)\end{cases}
$$

Let $T^{\prime}>0$ be small enough, so that $\left(h, u_{0}\right) \in W$. Let $z \in V$ be the solution of (10.3) and let $u \in \operatorname{MR}_{p}\left(0, T^{\prime} ; X, D\right)$ be the restriction of $z$ to the interval $\left(0, T^{\prime}\right)$. Then $u$ is a solution of (10.1).

### 10.2 Gradients of quadratic forms and identification of the trace space

Let $V$ and $H$ be two separable Hilbert spaces with inner products $\langle\cdot, \cdot\rangle_{V}$ and $\langle\cdot, \cdot\rangle_{H}$, respectively, and assume that $V$ is continuously and densely embedded into $H$. Consider the quadratic form $\mathscr{E}: V \rightarrow \mathbb{R}, \mathscr{E}(u)=\frac{1}{2}\|u\|_{V}^{2}$, and let $\nabla_{H} \mathscr{E}$ be the gradient of $\mathscr{E}$ with respect to the inner product $\langle\cdot, \cdot\rangle_{H}$. We recall from Exercise 8.2 that $\nabla_{H} \mathscr{E}$ is a linear operator and that the domain $D:=D\left(\nabla_{H} \mathscr{E}\right)$ is a Hilbert space for the inner product

$$
\langle u, v\rangle_{D}:=\left\langle\nabla_{H} \mathscr{E}(u), \nabla_{H} \mathscr{E}(v)\right\rangle_{H}+\langle u, v\rangle_{H} \quad(u, v \in D)
$$

One has

$$
D \subseteq V \subseteq H
$$

with continuous embeddings. Moreover, the space

$$
\operatorname{MR}_{2}(0, T ; H, D):=W^{1,2}(0, T ; H) \cap L^{2}(0, T ; D)
$$

is a Hilbert space for the inner product

$$
\langle u, v\rangle_{\mathrm{MR}_{2}}:=\int_{0}^{T}\langle u(t), v(t)\rangle_{D} d t+\int_{0}^{T}\langle\dot{u}(t), \dot{v}(t)\rangle_{H} d t .
$$

We omit the proof of the following lemma. It is very similar to the proof that $C^{1}([0, T])$ is dense in $W^{1, p}(0, T)$ (compare with Corollary G.13).

Lemma 10.3. The space $C^{1}([0, T] ; D)$ is dense in $\mathrm{MR}_{2}(0, T ; H, D)$.
Lemma 10.4. For every $u \in \operatorname{MR}_{2}(0, T ; H, D)$ one has $\|u\|_{V}^{2} \in W^{1,1}(0, T)$ and

$$
\frac{1}{2} \frac{d}{d t}\|u\|_{V}^{2}=\left\langle\nabla_{H} \mathscr{E}(u), \dot{u}\right\rangle_{H}
$$

the derivative on the left-hand side of this inequality being the weak derivative.
Proof. If $u \in C^{1}([0, T] ; D)$, then the function $\|u\|_{V}^{2}$ is continuously differentiable and, by definition of the gradient $\nabla_{H} \mathscr{E}$,

$$
\frac{d}{d t}\|u\|_{V}^{2}=\frac{d}{d t}\langle u, u\rangle_{V}=2\langle u, \dot{u}\rangle_{V}=2\left\langle\nabla_{H} \mathscr{E}(u), \dot{u}\right\rangle_{H}
$$

Let $u \in \operatorname{MR}_{2}(0, T ; H, D)$. By Lemma 10.3, there exists a sequence $\left(u_{n}\right) \subseteq$ $C^{1}([0, T] ; D)$ such that

$$
\begin{array}{ll}
u_{n} \rightarrow u & \text { in } L^{2}(0, T ; D), \\
\nabla_{H} \mathscr{E}\left(u_{n}\right) \rightarrow \nabla_{H} \mathscr{E}(u) & \text { in } L^{2}(0, T ; H), \text { and }  \tag{10.4}\\
\dot{u}_{n} \rightarrow \dot{u} & \text { in } L^{2}(0, T ; H) .
\end{array}
$$

By the above calculation and by the rule of integration by parts, for every $\varphi \in$ $C_{c}^{1}(0, T)$ and every $n$,

$$
\int_{0}^{T}\left\|u_{n}\right\|_{V}^{2} \dot{\varphi}=-2 \int_{0}^{T}\left\langle\nabla_{H} \mathscr{E}\left(u_{n}\right), \dot{u}_{n}\right\rangle_{H} \varphi
$$

The convergences in (10.4) allow us to pass to the limit on both sides of this equation. We thus obtain that for every $u \in \mathrm{MR}_{2}(0, T ; H, D)$ and every $\varphi \in C_{c}^{1}(0, T)$

$$
\int_{0}^{T}\|u\|_{V}^{2} \dot{\varphi}=-2 \int_{0}^{T}\left\langle\nabla_{H} \mathscr{E}(u), \dot{u}\right\rangle_{H} \varphi
$$

The claim follows from this equality, the definition of the Sobolev space $W^{1,1}(0,1)$ and the fact that, by the Cauchy-Schwarz inequality, $\left\langle\nabla_{H} \mathscr{E}_{\omega}(u), \dot{u}\right\rangle_{H} \in L^{1}(0, T)$.

Lemma 10.5. One has

$$
\mathrm{MR}_{2}(0, T ; H, D) \subseteq C([0, T] ; V)
$$

with continuous embedding.

Proof. For every $u \in C^{1}([0, T] ; D)$ we have, by Lemma 10.4 , and by continuity of the Sobolev embedding $W^{1,1}(0, T) \subseteq C([0, T])$ (applied to the function $\|u\|_{V}^{2}$ )

$$
\begin{aligned}
\|u\|_{C([0, T] ; V)}^{2} & =\sup _{t \in[0, T]}\|u(t)\|_{V}^{2}=\| \| u\left\|_{V}^{2}\right\|_{\infty} \\
& \leq C\left(\int_{0}^{T}\|u(t)\|_{V}^{2} d t+\int_{0}^{T}\left|\frac{d}{d t}\|u(t)\|_{V}^{2}\right| d t\right) \\
& \leq C\left(\int_{0}^{T}\|u(t)\|_{D}^{2} d t+2 \int_{0}^{T}\left\|\nabla_{H} \mathscr{E}(u(t))\right\|_{H}\|\dot{u}(t)\|_{H} d t\right) \\
& \leq C\|u\|_{\mathrm{MR}_{2}(0, T ; H, D)}^{2} .
\end{aligned}
$$

The constants $C$ here may change from line to line, they depend on the Sobolev embedding constant and on the constant of the embedding $D \hookrightarrow V$, but they do not depend on $u$. The claim thus follows from this estimate and Lemma 10.3.
Lemma 10.6. One has $\operatorname{Tr}_{2}(H, D)=V$.
Proof. The inclusion $\operatorname{Tr}_{2}(H, D) \subseteq V$ is a direct consequence of Lemma 10.5. For the other inclusion, recall that, by Theorem 6.1, for every $u_{0} \in V$ the gradient system

$$
\left\{\begin{array}{l}
\dot{u}+\nabla_{H} \mathscr{E}(u)=0, \\
u(0)=u_{0}
\end{array}\right.
$$

admits a (unique) solution $u \in W^{1,2}(0,1 ; H) \cap L^{\infty}(0,1 ; V)$. Since $\nabla_{H} \mathscr{E}(u)=$ $-\dot{u} \in L^{2}(0, T ; H)$, this solution belongs also to $L^{2}(0,1 ; D)$, and consequently to $\mathrm{MR}_{2}(0,1 ; H, D)$. Thus, by solving the gradient system, we have proved that for every $u_{0} \in V$ there exists $u \in \operatorname{MR}_{2}(0,1 ; H, D)$ such that $u(0)=u_{0}$. Since $u_{0} \in V$ was arbitrary, this proves the inclusion $V \subseteq \operatorname{Tr}_{2}(H, D)$.
Example 10.7. Let $V=H_{0}^{1}(0,1)$ and $H=L^{2}(0,1)$. We equip $L^{2}(0,1)$ with the usual $L^{2}$ inner product, and $H_{0}^{1}(0,1)$ with the inner product $\langle u, v\rangle_{H_{0}^{1}}=\int_{0}^{1} u^{\prime} v^{\prime}$. By Poincaré's inequality, this inner product is equivalent to the usual inner product which is induced from $H^{1}(0,1)$ (Exercise 4.7). The gradient of energy $\mathscr{E}$ : $H_{0}^{1}(0,1) \rightarrow \mathbb{R}, \mathscr{E}(u)=\frac{1}{2}\|u\|_{H_{0}^{1}}^{2}$ with respect to the usual $L^{2}$ inner product is the negative Dirichlet-Laplace operator on $L^{2}(0,1)$, that is

$$
\begin{aligned}
D\left(\nabla_{L^{2}} \mathscr{E}\right) & =H^{2}(0,1) \cap H_{0}^{1}(0,1), \\
\nabla_{L^{2}} \mathscr{E}(u) & =-u^{\prime \prime}=-{ }_{(0,1)}^{D} \Delta u
\end{aligned}
$$

see Lecture 3. By Lemma 10.6, we therefore find

$$
\begin{equation*}
\operatorname{Tr}_{2}\left(L^{2}(0,1), H^{2}(0,1) \cap H_{0}^{1}(0,1)\right)=H_{0}^{1}(0,1) \tag{10.5}
\end{equation*}
$$

Remark 10.8. a) More generally, if $\Omega \subseteq \mathbb{R}^{d}$ is a bounded, open set, and if ${ }_{\Omega}^{D} \Delta$ is the Dirichlet-Laplace operator on $L^{2}(\Omega)$ (see Lecture 4), then

$$
\operatorname{Tr}_{2}\left(L^{2}(\Omega), D\left({ }_{\Omega}^{D} \Delta\right)\right)=H_{0}^{1}(\Omega)
$$

One may argue as in the preceding example, by noting that the Poincaré inequality holds for every bounded open $\Omega \subseteq \mathbb{R}^{d}$ ([Adams (1975)], [Brézis (1992)]).
b) It is actually not necessary to appeal to Poincaré's inequality, for Lemma 10.6 remains true in a slightly more general setting. Assume that $\mathscr{E}: V \rightarrow \mathbb{R}$ is a continuous, $H$-elliptic, quadratic form. Still, by Exercise 8.2, the gradient $\nabla_{H} \mathscr{E}$ is a linear operator and the domain $D:=D\left(\nabla_{H} \mathscr{E}\right)$ is a Hilbert space. The equality

$$
\operatorname{Tr}_{2}(H, D)=V
$$

is still true in this situation (Exercise).
c) In particular, if $H=L^{2}(0,1)$ and $D=\left\{u \in H^{2}(0,1): u^{\prime}(0)=u^{\prime}(1)=0\right\}$, then

$$
\operatorname{Tr}_{2}(H, D)=H^{1}(0,1)
$$

This can be seen by considering the Neumann-Laplace operator on $L^{2}(0,1)$, for $D$ is just the domain of this operator, and $H^{1}(0,1)$ is the domain of the associated energy (see Exercise 4.3).

### 10.3 A semilinear diffusion equation

Let $f: \mathbb{R} \rightarrow \mathbb{R}$ be $k$ times continuously differentiable for some $k \geq 1$, and consider the semilinear diffusion equation

$$
\begin{cases}u_{t}-u_{x x}+f(u)=0 & \text { in }(0, \infty) \times(0,1)  \tag{10.6}\\ u(t, 0)=u(t, 1)=0 & \text { for } t \in(0, \infty) \\ u(0, x)=u_{0}(x) & \text { for } x \in(0,1)\end{cases}
$$

We rewrite the equation (10.6) as an abstract gradient system and we prove existence and regularity of a solution for that problem. We consider the energy $\mathscr{E}: H_{0}^{1}(0,1) \rightarrow \mathbb{R}$ given by

$$
\mathscr{E}(u)=\frac{1}{2} \int_{0}^{1} u_{x}^{2}+\int_{0}^{1} F(u),
$$

where $F(s)=\int_{0}^{s}(r) d r$. By Exercise 4.6, the energy $\mathscr{E}$ is continuously differentiable. Moreover, if $\nabla_{L^{2}} \mathscr{E}$ is the gradient with respect to the usual $L^{2}$ inner product, then

$$
\begin{aligned}
D\left(\nabla_{L^{2}} \mathscr{E}\right) & =H^{2}(0,1) \cap H_{0}^{1}(0,1), \\
\nabla_{L^{2}} \mathscr{E}(u) & =-{ }_{(0,1)}^{D} \Delta u+f(u) .
\end{aligned}
$$

Here ${ }_{(0,1)}^{D} \Delta$ is the Dirichlet-Laplace operator on $L^{2}(0,1)$ as it was defined in Lecture 3 , and $f$ is the composition operator $H_{0}^{1}(0,1) \rightarrow L^{2}(0,1), u \mapsto f(u)=f \circ u$; we use the same letter for the original function $f$ and any composition operator associated with this function since we think that there is no danger of confusion. In the following, we denote

$$
\begin{aligned}
H & :=L^{2}(0,1) \\
V & :=H_{0}^{1}(0,1), \text { and } \\
D & :=D\left(\nabla_{L^{2}} \mathscr{E}\right)=H^{2}(0,1) \cap H_{0}^{1}(0,1)
\end{aligned}
$$

Recall from Example 10.7 that

$$
\operatorname{Tr}_{2}\left(L^{2}(0,1), H^{2}(0,1) \cap H_{0}^{1}(0,1)\right)=H_{0}^{1}(0,1)
$$

Moreover, by Lemma 10.5, and by the Sobolev embedding $H_{0}^{1}(0,1) \subseteq C([0,1])$ (Theorem 5.11), we have the continuous embeddings

$$
\begin{equation*}
\operatorname{MR}_{2}(0, T ; H, D) \subseteq C\left([0, T] ; H_{0}^{1}(0,1)\right) \subseteq C([0, T] ; C([0,1])) \tag{10.7}
\end{equation*}
$$

Throughout this section, we deliberately use the identifications

$$
\begin{aligned}
& C([0, T] ; C([0,1]))=C([0, T] \times[0,1]) \text { and } \\
& C((0, T] ; C([0,1]))=C((0, T] \times[0,1])
\end{aligned}
$$

and their consequences. It follows from these identifications that we may identify a solution $u$ of the gradient system (10.8) below with a real-valued function of two variables (time and space variable), or, vice versa, we may identify a solution of the semilinear diffusion equation (10.6) with an element of a Bochner or BochnerSobolev space. The aim is of course to show that every solution of the gradient system (10.8) below gives rise to a solution of the semilinear diffusion equation (10.6), and vice versa. However, here we concentrate on the abstract gradient system.

Theorem 10.9. For every $u_{0} \in H_{0}^{1}(0,1)$ the problem

$$
\left\{\begin{array}{l}
\dot{u}-\underset{(0,1)}{D} \Delta u+f(u)=0,  \tag{10.8}\\
u(0)=u_{0},
\end{array}\right.
$$

admits a unique local solution

$$
\begin{aligned}
& u \in \operatorname{MR}_{2}(0, T ; H, D) \text { satisfying in addition } \\
& u \in W_{l o c}^{k+1,2}\left((0, T] ; L^{2}(0,1)\right) \cap W_{l o c}^{k, 2}\left((0, T] ; H^{2}(0,1)\right) \cap C^{k}\left((0, T] ; H_{0}^{1}(0,1)\right) .
\end{aligned}
$$

This theorem is an application of the Theorems 10.1 and 9.1. In order to be able to apply both theorems, we need the following lemmas.

Lemma 10.10. a) The gradient $\nabla_{L^{2} \mathscr{E}}: D \rightarrow H$ is of class $C^{k}$.
b) The composition operator $\mathrm{MR}_{2}(0, T ; H, D) \rightarrow L^{2}(0, T ; H), u \mapsto \nabla_{L^{2}} \mathscr{E}(u)$ is everywhere defined and of class $C^{k}$. The derivative $\left(\nabla_{L^{2}} \mathscr{E}\right)^{\prime}(u)$ : $\mathrm{MR}_{2}(0, T ; H, D) \rightarrow L^{2}(0, T ; H)$ at a point $u$ is given by

$$
\left(\nabla_{L^{2}} \mathscr{E}\right)^{\prime}(u) v=-\underset{(0,1)}{D} \Delta v+f^{\prime}(u) v,
$$

where $f^{\prime}(u) v$ stands for the pointwise multplication of the composite function $f^{\prime}(u)=f^{\prime} \circ u$ and the function $v$.
Proof. The Dirichlet-Laplace operator is continuous and linear from $D$ into $H$. Hence, by Exercise 9.3, it suffices to study only the composition operator (Nemytski operator) associated with the function $f$.
(a) We may factorize the composition operator associated with $f$ in the following way:

where $i$ stands for the natural embeddings. The natural embedding $D \subseteq C([0,1])$ is continuous by the Sobolev embedding theorem (Theorem 5.11), and the natural embedding $C([0,1]) \subseteq H$ is continuous as a consequence of Hölder's inequality. With this factorization, since the embeddings $i$ are linear (and therefore of class $C^{\infty}$ ), and since the composition operator associated with $f$ leaves the $C([0,1])$ invariant, the claim follows from Exercise 9.2.
(b) We may factorize the composition operator associated with $f$ in the following way:

where $i$ stands for the natural embeddings and $b$ for the natural bijections. In the first embedding on the left, we used (10.7). With the above factorization, since the operators $i$ and $b$ are linear, and since the composition operator associated with $f$ leaves the $C([0,1])$ invariant, the claim follows from Exercise 9.2.
Lemma 10.11. For every $T>0$, every $c \in L^{2}\left(0, T ; L^{2}(0,1)\right)$, every $v_{0} \in H_{0}^{1}(0,1)$ and every $g \in L^{2}\left(0, T ; L^{2}(0,1)\right)$ the problem

$$
\left\{\begin{array}{l}
\dot{v}-{ }_{(0,1)}^{D} \Delta v+c v=g,  \tag{10.9}\\
v(0)=v_{0}
\end{array}\right.
$$

admits a unique solution $v \in \operatorname{MR}_{2}(0, T ; H, D)$.
The first line in (10.9) means $\dot{v}(t)-{ }_{(0,1)}^{D} \Delta v(t)+c(t) v(t)=g(t)$ for almost every $t \in(0, T)$ (equality in $L^{2}(0,1)$ ), and the term $c(t) v(t)$ stands for the pointwise multiplication of the functions $c(t) \in L^{2}(0,1)$ and $v(t) \in H_{0}^{1}(0,1) \subseteq C([0,1])$. Note that, by the embedding (10.7), for every $v \in \operatorname{MR}_{2}(0, T ; H, D)$ and every $c \in L^{2}\left(0, T ; L^{2}(0,1)\right)$ one has $c v \in L^{2}\left(0, T ; L^{2}(0,1)\right)$. In the proof of Lemma 10.11 we need the following variant of Gronwall's lemma.

Lemma 10.12 (Gronwall). Let $\lambda:[a, b] \rightarrow \mathbb{R}$ be integrable, and let $\varphi:[a, b] \rightarrow \mathbb{R}$ be a continuous functions on the bounded interval $[a, b]$. Assume that, for some $C \geq 0$ and every $t \in[a, b]$,

$$
\varphi(t) \leq C+\int_{a}^{t} \lambda(s) \varphi(s) d s
$$

Then, for every $t \in[a, b]$,

$$
\varphi(t) \leq C e^{\Lambda(t)}
$$

where $\Lambda(t)=\int_{0}^{t} \lambda(s) d s$.
Proof (of Lemma 10.11). Fix $T>0$. It suffices to show that for every $c \in$ $L^{2}\left(0, T ; L^{2}(0,1)\right)$ the linear operator

$$
\begin{aligned}
G_{c}: \mathrm{MR}_{2}(0, T ; H, D) & \rightarrow L^{2}\left(0, T ; L^{2}(0,1)\right) \times H_{0}^{1}(0,1), \\
v & \mapsto\left(\dot{v}-{ }_{(0,1)}^{D} \Delta v+c v, v(0)\right)
\end{aligned}
$$

is an isomorphism. Consider the set

$$
C:=\left\{c \in L^{2}\left(0, T ; L^{2}(0,1)\right): G_{c} \text { is an isomorphism }\right\} .
$$

For every $c_{1}, c_{2} \in L^{2}\left(0, T ; L^{2}(0,1)\right)$ one has

$$
\begin{aligned}
\left\|G_{c_{1}}-G_{c_{2}}\right\|_{\mathscr{L}} & =\sup _{\|u\|_{\mathrm{MR}_{2}} \leq 1}\left\|G_{c_{1}} u-G_{c_{2}} u\right\|_{L^{2}\left(0, T ; L^{2}\right) \times H_{0}^{1}} \\
& =\sup _{\|u\|_{\mathrm{MR}_{2}} \leq 1}\left\|\left(c_{1}-c_{2}\right) u\right\|_{L^{2}\left(0, T ; L^{2}\right)} \\
& \leq \sup _{\|u\|_{\mathrm{MR}_{2} \leq 1} \leq c_{1}-c_{2}\left\|_{L^{2}\left(0, T ; L^{2}\right)}\right\| u \|_{L^{\infty}\left(0, T ; L^{\infty}\right)}} \\
& \leq C\left\|c_{1}-c_{2}\right\|_{L^{2}\left(0, T ; L^{2}\right)}
\end{aligned}
$$

where $C$ is the constant of the embedding (10.7). Hence, the mapping $c \mapsto G_{c}$ is continuous from the space $L^{2}\left(0, T ; L^{2}(0,1)\right)$ into the space of continuous linear operators from $\mathrm{MR}_{2}(0, T ; H, D)$ into $L^{2}\left(0, T ; L^{2}(0,1)\right) \times H_{0}^{1}(0,1)$. Since the isomorphisms form an open subset in the set of all continuous, linear operators, we see that $C$ is open.

Next, we remark that for $c \equiv 0$, the problem (10.9) is nothing else than equation (7.11). By Theorem 7.1, for every $g \in L^{2}\left(0, T ; L^{2}(0,1)\right)$ and every $u_{0} \in H_{0}^{1}(0,1)$
the problem (10.9) (with $c \equiv 0$ ) admits a unique solution $v \in W^{1,2}\left(0, T ; L^{2}(0,1)\right) \cap$ $L^{\infty}\left(0, T ; H_{0}^{1}(0,1)\right)$. It follows from equation (10.9) that this solution satisfies ${ }_{(0,1)}^{D} \Delta v \in L^{2}(0, T ; H)$, and hence $v \in \operatorname{MR}_{2}(0, T ; H, D)$. Hence, $G_{0}$ is continuous, linear and bijective. By the bounded inverse theorem, the operator $G_{0}$ is an isomorphism. Hence, the set $C$ is nonempty.

Finally, let $c \in C$. Let $g \in L^{2}\left(0, T ; L^{2}(0,1)\right), v_{0} \in H_{0}^{1}(0,1)$, and let $v \in$ $\mathrm{MR}_{2}(0, T ; H, D)$ be the unique solution of the problem (10.9). By Lemma 10.4, the function $t \mapsto\|v(t)\|_{H_{0}^{1}}^{2}$ belongs to $W^{1,1}(0, T)$ and

$$
\frac{1}{2} \frac{d}{d t}\|v\|_{H_{0}^{1}}^{2}=-\left\langle{ }_{(0,1)}^{D} \Delta v, \dot{v}\right\rangle_{L^{2}} .
$$

Hence, if we multiply equation (10.9) by $\dot{v}$ with respect to the $L^{2}$ inner product and integrate the result over $(0, t)$ (with $t \in(0, T)$ ), then we obtain

$$
\begin{aligned}
\int_{0}^{t}\|\dot{v}\|_{L^{2}}^{2}+\frac{1}{2}\|v(t)\|_{H_{0}^{1}}^{2} & =\frac{1}{2}\left\|v_{0}\right\|_{H_{0}^{1}}^{2}+\int_{0}^{t}\langle g-c v, \dot{v}\rangle_{L^{2}} \\
& \leq \frac{1}{2}\left\|v_{0}\right\|_{H_{0}^{1}}^{2}+\int_{0}^{t}\|g\|_{L^{2}}^{2}+\int_{0}^{t}\|c\|_{L^{2}}^{2}\|v\|_{L^{\infty}}^{2}+\frac{1}{2} \int_{0}^{t}\|\dot{v}\|_{L^{2}}^{2}
\end{aligned}
$$

or

$$
\begin{equation*}
\frac{1}{2} \int_{0}^{t}\|\dot{v}\|_{L^{2}}^{2}+\frac{1}{2}\|v(t)\|_{H_{0}^{1}}^{2} \leq \frac{1}{2}\left\|v_{0}\right\|_{H_{0}^{1}}^{2}+\int_{0}^{T}\|g\|_{L^{2}}^{2}+\int_{0}^{t}\|c\|_{L^{2}}^{2}\|v\|_{H_{0}^{1}}^{2} \tag{10.10}
\end{equation*}
$$

Here, we have also applied the Sobolev embedding theorem (Theorem 5.11). By the variant of Gronwall's lemma (Lemma 10.12), the preceding inequality implies that

$$
\frac{1}{2}\|v(t)\|_{H_{0}^{1}}^{2} \leq\left(\frac{1}{2}\left\|v_{0}\right\|_{H_{0}^{1}}^{2}+\int_{0}^{T}\|g\|_{L^{2}}^{2}\right) e^{2 \int_{0}^{t}\|c\|_{L^{2}}^{2} \quad \text { for every } t \in(0, T) . . . ~}
$$

When we insert this estimate into (10.10), we obtain that for every $t \in(0, T)$

$$
\frac{1}{2} \int_{0}^{t}\|\dot{v}\|_{L^{2}}^{2}+\frac{1}{2}\|v(t)\|_{H_{0}^{1}}^{2} \leq\left(\frac{1}{2}\left\|v_{0}\right\|_{H_{0}^{1}}^{2}+\int_{0}^{T}\|g\|_{L^{2}}^{2} e^{2 \int_{0}^{t}\|c\|_{L^{2}}^{2}}\right.
$$

Hence,

$$
\|\dot{v}\|_{L^{2}\left(0, T ; L^{2}\right)}^{2}+\frac{1}{2}\|v\|_{L^{\infty}\left(0, T ; H_{0}^{1}\right)}^{2} \leq 2 e^{2\|c\|_{L^{2}\left(0, T ; L^{2}\right)}^{2}}\left(\frac{1}{2}\left\|v_{0}\right\|_{H_{0}^{1}}^{2}+\int_{0}^{T}\|g\|_{L^{2}}^{2}\right)
$$

From equation (10.9) and the Sobolev embedding $H_{0}^{1}(0,1) \subseteq C([0,1])$ we obtain

$$
\left\|_{(0,1)}^{D} \Delta v\right\|_{L^{2}\left(0, T ; L^{2}\right)} \leq\|\dot{v}\|_{L^{2}\left(0, T ; L^{2}\right)}+\|c\|_{L^{2}\left(0, T ; L^{2}\right)}\|v\|_{L^{\infty}\left(0, T ; H_{0}^{1}\right)}+\|g\|_{L^{2}\left(0, T ; L^{2}\right)} .
$$

This inequality and the preceding estimate finally imply that for every $R \geq 0$ there exists $K_{R} \geq 0$ such that whenever $\|c\|_{L^{2}\left(0, T ; L^{2}\right)} \leq R$, then

$$
\|v\|_{\mathrm{MR}_{2}}^{2}=\left\|G_{c}^{-1}\left(g, v_{0}\right)\right\|_{\mathrm{MR}_{2}}^{2} \leq K_{R}\left(\|g\|_{L^{2}\left(0, T ; L^{2}\right)}^{2}+\left\|v_{0}\right\|_{H_{0}^{1}}^{2}\right)
$$

In other words, $G_{c}^{-1}$ is uniformly bounded whenever $c$ varies in bounded subsets of $C$. This implies that $C$ must be closed.

Since $C \subseteq L^{2}\left(0, T ; L^{2}(0,1)\right)$ is open, closed and nonempty, and since the space $L^{2}\left(0, T ; L^{2}(0,1)\right)$ is connected, we obtain that $C=L^{2}\left(0, T ; L^{2}(0,1)\right)$. The claim follows from the definition of $C$ and since $T>0$ was arbitrary.

Proof (of Theorem 10.9). Existence. Let $u_{0} \in H_{0}^{1}(0,1)$. Since, by Lemma 10.6, $V=$ $H_{0}^{1}(0,1)=\operatorname{Tr}_{2}(H, D)$, we can choose $\bar{u} \in \operatorname{MR}_{2}(0,1 ; H, D)$ such that $\bar{u}(0)=u_{0}$. By (10.7), $\bar{u} \in C([0,1] ; C([0,1]))$, and therefore

$$
f^{\prime}(\bar{u}) \in C([0,1] ; C([0,1])) \subseteq L^{2}\left(0,1 ; L^{2}(0,1)\right)
$$

since $f^{\prime}$ is continuous. By Lemma 10.11, for every $g \in L^{2}\left(0,1 ; L^{2}(0,1)\right)$ and every $v_{0} \in H_{0}^{1}(0,1)$, the problem

$$
\left\{\begin{array}{l}
\dot{v}-{ }_{(0,1)}^{D} \Delta v+f^{\prime}(\bar{u}) v=g, \\
v(0)=v_{0}
\end{array}\right.
$$

admits a unique solution $v \in \operatorname{MR}_{2}(0,1 ; H, D)$. Since, by Lemma 10.10, $\left(\nabla_{L^{2}} \mathscr{E}\right)^{\prime}(\bar{u}) v=-{ }_{(0,1)}^{D} \Delta v+f^{\prime}(\bar{u}) v$, we may therefore apply Theorem 10.1 and we obtain a local solution $u \in \mathrm{MR}_{2}\left(0, T^{\prime} ; H, D\right)$ of the problem (10.8). Again, for this solution we have

$$
f^{\prime}(u) \in C([0, T] ; C([0,1])) \subseteq L^{2}\left(0, T ; L^{2}(0,1)\right) \quad\left(T \in\left(0, T^{\prime}\right)\right)
$$

since $f^{\prime}$ is continuous. By Lemma 10.11, for every $g \in L^{2}\left(0, T ; L^{2}(0,1)\right)$ and every $v_{0} \in H_{0}^{1}(0,1)$, the problem

$$
\left\{\begin{array}{l}
\dot{v}-{ }_{(0,1)}^{D} \Delta v+f^{\prime}(u) v=g, \\
v(0)=v_{0}
\end{array}\right.
$$

admits a unique solution $v \in \mathrm{MR}_{2}(0, T ; H, D)$. By Theorem 9.1 and Lemma 10.10, we obtain that the solution satisfies the stated regularity.
Uniqueness. See Exercises.

### 10.4 Exercises

10.1. Let $V$ and $H$ be two Hilbert spaces such that $V$ is densely and continuously embedded into $H$. Let $\mathscr{E}: V \rightarrow \mathbb{R}$ be a continuous, $H$-elliptic, quadratic form. By Exercise 8.2, the gradient $\nabla_{H} \mathscr{E}$ is a linear operator and the domain $D:=D\left(\nabla_{H} \mathscr{E}\right)$ is a Hilbert space. Show the equality

$$
\operatorname{Tr}_{2}(H, D)=V
$$

## 10.2 (Gronwall's lemma). Prove the Lemma 10.12.

10.3. Prove the uniqueness of local solutions in Theorem 10.9.

Hint: Assume that $u \in \operatorname{MR}_{2}(0, T ; H, D)$ and $v \in \operatorname{MR}_{2}\left(0, T^{\prime} ; H, D\right)$ are two local solutions. By using the embedding (10.7) and the fact that the function $f$ is locally Lipschitz continuous, establish a differential inequality for the function $\|u(\cdot)-v(\cdot)\|_{L^{2}}^{2}$ ( $0 \leq t \leq \min \left\{T, T^{\prime}\right\}$ ).
10.4. Consider the semilinear diffusion equation

$$
\begin{cases}u_{t}-u_{x x}-u^{3}=0 & \text { in }(0, \infty) \times(0,1)  \tag{10.11}\\ u(t, 0)=u(t, 1)=0 & \text { for } t \in(0, \infty) \\ u(0, x)=u_{0}(x) & \text { for } x \in(0,1)\end{cases}
$$

This equation can be rewritten as an abstract gradient system for the energy $\mathscr{E}$ : $H_{0}^{1}(0,1) \rightarrow \mathbb{R}$ given by

$$
\mathscr{E}(u)=\frac{1}{2} \int_{0}^{1} u_{x}^{2}-\frac{1}{4} \int_{0}^{1} u^{4}
$$

a) Show that the energy $\mathscr{E}$ is not $L^{2}(0,1)$-elliptic.
b) Show that for every $u_{0} \in H^{1}(0,1)$ satisfying $\mathscr{E}\left(u_{0}\right)<0$ the unique local solution of (10.8) (which exists by Theorem 10.9) can not be extended to a global solution defined on $(0, \infty)$.
Hint. Multiply the gradient system (10.8) by $u$ with respect to the usual $L^{2}$ inner product. By using that $\mathscr{E}(u)$ is decreasing with respect to the time variable, establish a differential inequality for $\|u\|_{L^{2}}^{2}$ and deduce that there must be blow-up in finite time, that is, that there exists $T<\infty$ such that $\lim _{t \rightarrow T-}\|u(t)\|_{L^{2}}^{2}=\infty$.

## Lecture 11

## Asymptotic behaviour of solutions of gradient systems

The gradient system

$$
\begin{equation*}
\dot{u}+\nabla_{g} \mathscr{E}(u)=0 \tag{11.1}
\end{equation*}
$$

is a prototype example of a dissipative system, that is, by our loose definition from Lecture 1, a system which admits an energy function. An energy function is, by definition, a function which is decreasing along every solution. In the case of a gradient system in finite-dimensional space, by Lemma 2.5 , the function $\mathscr{E}$ itself is an energy, and it satsfies the stronger property that if it is constant along a solution then that solution must already be constant. For gradient systems in infinite-dimensional spaces we did not prove an analogous statement, but the energy inequalities in Theorems 6.1 and 8.1 and the energy equality in Lemma 10.5 indicate that some gradient systems in infinite-dimensional spaces are dissipative, too. Actually, we are not aware of a gradient system which is not dissipative. Concerning the meaning of "dissipative" or "to dissipate", let us recall the footnote from page 11:
to dissipate (lat.: dissipare): to cause to lose energy (such as heat) irreversibly. Or: to spend or expend intemperately or wastefully. Or: to attenuate to or almost to the point of disappearing.

In this explanation, the last point is of particular interest in connection with this lecture in which we study the asymptotic behaviour of solutions of gradient systems. What does it mean, "to attenuate to or almost to the point of disappearing"?

First, it should be noted that, by Lemma 1.6, every limit point of a global solution of a Euclidean gradient system is an equilibrium point, that is, a point in which the gradient vanishes. It is not too difficult to see that this result is also true for general gradient systems in finite-dimensional spaces, and for gradient systems in infinitedimensional spaces if some energy inequality is true along the solution (Lemma 11.3). A basic problem, however, is to know whether every global "bounded" solutions has only one limit point, that is, whether it converges to a single equilibrium point = "the point of disappearing". Moreover, if a global, bounded solution converges to a single equilibrium point, can we determine the rate of convergence? Such questions are studied in this and the following lecture. The problem of convergence
to equilibrium admits a sort of dichotomy: there is a relatively simple convergence result which applies in situations where we know something about the topology of the set of equilibrium points, in particular, that the set of equilibrium points is discrete (Theorem 11.4). However, if we know nothing about the set of equilibrium points, or if we know that there is a continuum of equilibrium points, then the problem of convergence becomes very quickly quite involved. In general, we must say that the above description of dissipativity does not coincide with the reality of gradient systems: there exists an example of a bounded solution of a gradient system which does not converge! The existence of such an example was already suggested in words by Curry in 1944; a concrete example was then given in 1982 by Palis and de Melo. We describe this example of a nonconvergent soltuion at the end of this lecture.

### 11.1 The $\omega$-limit set of a continuous function on $\mathbb{R}_{+}$

Let $(M, d)$ be a metric space, and let $u: \mathbb{R}_{+} \rightarrow M$ be a continuous function. The set

$$
\omega(u)=\left\{\varphi \in M: \text { there exists an unbounded }\left(t_{n}\right) \subseteq \mathbb{R}_{+} \text {such that } \lim _{n \rightarrow \infty} u\left(t_{n}\right)=\varphi\right\} .
$$

is called the $\omega$-limit set ${ }^{1}$ of $u$. Clearly, in this definition, by passing to appropriate subsequences, one may replace "unbounded $\left(t_{n}\right)$ " by "unbounded and increasing $\left(t_{n}\right)$ ". Alternatively, one may define the $\omega$-limit set in the following way.

## Lemma 11.1. One has

$$
\omega(u)=\bigcap_{t \geq 0} \overline{\{u(s): s \geq t\}}
$$

Proof. Let $\varphi \in \omega(u)$. By definition, there exists an unbounded sequence $\left(t_{n}\right) \subseteq \mathbb{R}_{+}$ such that $\lim _{n \rightarrow \infty} u\left(t_{n}\right)=\varphi$. Since the sequence $\left(t_{n}\right)$ is unbounded, this implies $\varphi \in$ $\overline{\{u(s): s \geq t\}}$ for every $t \geq 0$, and hence $\varphi \in \bigcap_{t \geq 0} \overline{\{u(s): s \geq t\}}$.

Conversely, let $\varphi \in \bigcap_{t \geq 0} \overline{\{u(s): s \geq t\}}$. Then, for every $n \in \mathbb{N}, \varphi \in \overline{\{u(s): s \geq n\}}$. In particular, for every $n \in \mathbb{N}$, there exists $t_{n} \geq n$ such that $d\left(u\left(t_{n}\right), \varphi\right) \leq \frac{1}{n}$. We obtain therefore an unbounded sequence $\left(t_{n}\right) \subseteq \mathbb{R}_{+}$such that $\lim _{n \rightarrow \infty} u\left(t_{n}\right)=\varphi$. By definition, $\varphi \in \omega(u)$.

Lemma 11.2. Assume that $u: \mathbb{R}_{+} \rightarrow M$ has relatively compact range. Then:
a) The $\omega$-limit set is non-empty, compact and connected.
b) If one denotes $\operatorname{dist}(u(t), \omega(u)):=\inf \{d(u(t), \varphi): \varphi \in \omega(u)\}$, then one has $\lim _{t \rightarrow \infty} \operatorname{dist}(u(t), \omega(u))=0$.

[^10]c) The limit $\lim _{t \rightarrow \infty} u(t)$ exists if and only if $\omega(u)$ contains exactly one element.

Proof. (a) Since $u$ is continuous and has relatively compact range, for every $t \geq 0$ the set $K_{t}:=\overline{\{u(s): s \geq t\}}$ is non-empty, compact and connected. Moreover, the family $\left(K_{t}\right)$ is decreasing, that is, $K_{t} \subseteq K_{s}$ for $t \geq s$. Then the characterization from Lemma 11.1 implies that $\omega(u)$ is non-empty, compact and connected, too.
(b) Assume that the claim is not true. Then there exists an unbounded sequence $\left(t_{n}\right) \subseteq \mathbb{R}_{+}$such that $\inf \left\{\operatorname{dist}\left(u\left(t_{n}\right), \omega(u)\right): n \geq 1\right\}>0$. Since $u$ has relatively compact range, we can extract an unbounded subsequence $\left(t_{n_{k}}\right)$ of $\left(t_{n}\right)$, and find $\varphi \in V$ such that $\lim _{k \rightarrow \infty} u\left(t_{n_{k}}\right)=\varphi$. By definition, $\varphi \in \omega(u)$. This implies $\lim _{k \rightarrow \infty} \operatorname{dist}\left(u\left(t_{n_{k}}\right), \omega(u)\right)=0$, a contradiction.
(c) If the limit $\lim _{t \rightarrow \infty} u(t)=: \varphi$ exists, then $\omega(u)=\{\varphi\}$; this follows from the definition of the $\omega$-limit set. On the other hand, if $\omega(u)=\{\varphi\}$, then (b) implies that $\lim _{t \rightarrow \infty} u(t)=\varphi$.

### 11.2 A stabilisation result for global solutions of gradient systems

Let $U$ be an open subset of a Banach space $V$ which embeds densely and continuously into a Hilbert space $H$. Let $\mathscr{E}: U \rightarrow \mathbb{R}$ be a continuously differentiable function, and let $g: U \rightarrow \operatorname{Inner}(H)$ be a metric. We assume that there exists a constant $c \geq 0$ such that

$$
\begin{equation*}
\|v\|_{H} \leq c\|v\|_{g(u)} \quad \text { for every } v \in H, u \in U \tag{11.2}
\end{equation*}
$$

Lemma 11.3. Let $u \in W_{\text {loc }}^{1,2}\left(\mathbb{R}_{+} ; H\right) \cap C\left(\mathbb{R}_{+} ; V\right)$ be a global solution of the gradient system (11.1). Assume that u has relatively compact range in $U$, and that the energy inequality

$$
\begin{equation*}
\int_{0}^{t}\|\dot{u}(s)\|_{g(u(s))}^{2} d s+\mathscr{E}(u(t)) \leq \mathscr{E}(u(0)) \tag{11.3}
\end{equation*}
$$

holds for every $t \in \mathbb{R}_{+}$. Then every $\omega$-limit point of $u$ is an equilibrium point of $\mathscr{E}$, that is, every element $\varphi \in \omega(u)$ satisfies $\varphi \in D\left(\nabla_{g} \mathscr{E}\right)$ and $\nabla_{g} \mathscr{E}(\varphi)=0$.

Proof. Since $u$ has relatively compact range in $U$, and since $\mathscr{E}$ is continuous on $U$, the composite function $\mathscr{E}(u)$ is necessarily bounded. Then the energy inequality (11.3) implies that the integral $\int_{0}^{\infty}\|\dot{u}(s)\|_{g(u(s))}^{2} d s$ is finite. By assumption (11.2), the integral $\int_{0}^{\infty}\|\dot{u}(s)\|_{H}^{2}$ is thus finite, too.

Let $\varphi \in \omega(u)$. Then there exists an unbounded increasing sequence $\left(t_{n}\right) \subseteq \mathbb{R}_{+}$ such that $\lim _{n \rightarrow \infty}\left\|u\left(t_{n}\right)-\varphi\right\|_{V}=0$. Since $V$ is continuously embedded into $H$, and by Hölder's inequality, we obtain for every $s \in[0,1]$

$$
\begin{aligned}
\limsup _{n \rightarrow \infty}\left\|u\left(t_{n}+s\right)-\varphi\right\|_{H} & \leq \limsup _{n \rightarrow \infty}\left(\left\|u\left(t_{n}\right)-\varphi\right\|_{H}+\int_{t_{n}}^{t_{n}+s}\|\dot{u}(r)\|_{H} d r\right) \\
& \leq \limsup _{n \rightarrow \infty}\left(\int_{t_{n}}^{t_{n}+s}\|\dot{u}(r)\|_{H}^{2} d r\right)^{\frac{1}{2}}=0
\end{aligned}
$$

Using again the relative compactness of the range of $u$ and a sub-subsequence argument, this implies that

$$
\lim _{n \rightarrow \infty}\left\|u\left(t_{n}+s\right)-\varphi\right\|_{V}=0 \quad \text { for every } s \in[0,1]
$$

By continuity of $\mathscr{E}^{\prime}$,

$$
\lim _{n \rightarrow \infty}\left\|\mathscr{E}^{\prime}\left(u\left(t_{n}+s\right)\right)-\mathscr{E}^{\prime}(\varphi)\right\|_{V^{\prime}}=0 \text { for every } s \in[0,1]
$$

Therefore, by the dominated convergence theorem, by definition of the gradient, and since $u$ is a solution of the gradient system (11.1), for every $v \in V$,

$$
\begin{aligned}
\left|\left\langle\mathscr{E}^{\prime}(\varphi), v\right\rangle\right| & =\left|\int_{0}^{1}\left\langle\mathscr{E}^{\prime}(\varphi), v\right\rangle d s\right| \\
& =\lim _{n \rightarrow \infty}\left|\int_{0}^{1}\left\langle\mathscr{E}^{\prime}\left(u\left(t_{n}+s\right)\right), v\right\rangle d s\right| \\
& =\lim _{n \rightarrow \infty}\left|\int_{0}^{1}\left\langle\dot{u}\left(t_{n}+s\right), v\right\rangle_{g\left(u\left(t_{n}+s\right)\right)} d s\right|
\end{aligned}
$$

By continuity of the metric $g$, and since $u$ has relatively compact range in $U$, for every $v, w \in H$ the set $\left\{\left|\langle v, w\rangle_{g(u(t))}\right|: t \in \mathbb{R}_{+}\right\}$is bounded. Hence, by the uniform boundedness principle, there exists a constant $C \geq 0$ such that for every $v \in H$ and every $t \in \mathbb{R}_{+}$one has $\|v\|_{g(u(t))}^{2} \leq C\|v\|_{H}^{2}$. Hence, for every $v \in V$,

$$
\begin{aligned}
\left|\left\langle\mathscr{E}^{\prime}(\varphi), v\right\rangle\right| & \leq \lim _{n \rightarrow \infty}\left(\int_{0}^{1}\left\|\dot{u}\left(t_{n}+s\right)\right\|_{g\left(u\left(t_{n}+s\right)\right)}^{2} d s\right)^{\frac{1}{2}}\left(\int_{0}^{1}\|v\|_{g\left(u\left(t_{n}+s\right)\right)}^{2}\right)^{\frac{1}{2}} \\
& \leq \lim _{n \rightarrow \infty}\left(\int_{0}^{1}\left\|\dot{u}\left(t_{n}+s\right)\right\|_{g\left(u\left(t_{n}+s\right)\right)}^{2} d s\right)^{\frac{1}{2}} \sqrt{C}\|v\|_{H} \\
& =0 .
\end{aligned}
$$

This yields $\mathscr{E}^{\prime}(\varphi)=0$ and the claim.
A topological Hausdorff space $S$ is discrete if the topology on $S$ is the discrete topology, that is, every subset of $S$ is open and closed. In a discrete space the only connected subsets are the subsets which contain at most one element. This observation leads to the following stabilisation result for global solutions of gradient systems.

Theorem 11.4. Assume that the set of equilibrium points $S:=\left\{\varphi \in D\left(\nabla_{g} \mathscr{E}\right)\right.$ : $\left.\nabla_{g} \mathscr{E}(\varphi)=0\right\} \subseteq U$ is discrete. Then every global solution $u$ of (11.1) with relatively compact range in $U$ and satisfying the energy inequality (11.3) converges to an equilibrium, that is, there exists $\varphi \in S$ such that $\lim _{t \rightarrow \infty} u(t)=\varphi$ in $V$.

Proof. By Lemma 11.3, the $\omega$-limit set of every continuous, global solution having relatively compact range in $U$ is contained in the set of equilibrium points. By Lemma 11.2 (a), the $\omega$-limit set is non-empty and connected. Since, by assumption,
the set of equilibrium points is discrete, we deduce that the $\omega$-limit set contains exactly one point. The claim now follows from Lemma 11.2 (c).

### 11.3 Nonstabilisation results for global solutions of gradient systems

Theorem 11.4 provides a very useful criterion for stabilization of global solutions having relatively compact range. Exercise 11.3 contains a nontrivial example of a semilinear diffusion equation where the set of equilibrium points is discrete. However, one can easily imagine situations where the set of equilibrium points is nondiscrete: think of an energy which is constant on a nonempty open set or look at Exercise 11.4. In those situations, can we still expect convergence to a single equilibrium point? Is convergence a consequence of the gradient structure and/or dissipativity? The following suggests that the answer is no, although there is no explicit example.

> The following example shows that we cannot expect a better result without further restrictions on $\mathscr{E}$. Let $\mathscr{E}(x, y)=0$ on the unit circle and $\mathscr{E}(x, y)>0$ elsewhere. Outside the unit circle let the surface have a spiral gully making infinitely many turns about the circle. Then the path $C$ will evidently follow the gully and have all points of the unit circle as limit points. ${ }^{2}$

Haskell B. Curry
It is not clear whether Curry had a concrete energy $\mathscr{E}$ in mind. In fact, it took quite a while until in [Palis and de Melo (1982), page 14] the following $C^{\infty}$ "Mexicanhat" function

$$
\mathscr{E}(r \cos \theta, r \sin \theta)= \begin{cases}\exp \left(\frac{1}{r^{2}-1}\right) & \text { if } r<1 \\ 0 & \text { if } r=1 \\ \exp \left(-\frac{1}{r^{2}-1}\right) \sin \left(\frac{1}{r-1}-\theta\right) & \text { if } r>1\end{cases}
$$

appeared. Palis and de Melo showed by some abstract arguments that the Euclidean gradient system associated with this energy possesses a global, bounded solution which has the whole unit circle as $\omega$-limit set. For curiosity we note that in their example the solution stays outside the unit circle and the particular shape of $\mathscr{E}$ inside the unit circle is not important. In fact, there are many ways how to extend $\mathscr{E}$ to a global $C^{\infty}$ function inside the unit disk, and the particular definition of Palis and de Melo is perhaps only motivated by the desire to obtain a "Mexican hat".

We do not repeat the analysis of Palis and de Melo, but rather consider the following energy

[^11]

Fig. 11.1 The graph of the "mexican hat" of Palis \& de Melo

$$
\mathscr{E}(r \cos \theta, r \sin \theta)= \begin{cases}\exp \left(\frac{1}{r^{2}-1}\right) & \text { if } r<1 \\ 0 & \text { if } r=1 \\ \exp \left(\frac{-1}{r^{2}-1}\right)\left[1+\frac{4 r^{4}}{4 r^{4}+\left(1-r^{2}\right)^{4}} \sin \left(\frac{1}{r^{2}-1}-\theta\right)\right] \text { if } r>1\end{cases}
$$

which was essentially proposed by Absil, Mahony and Andrews in [Absil et al. (2005)]. It looks more complicated than the energy of Palis and de Melo, but it is easier to show that this energy gives an example of nonstabilization. Note that the Euclidean gradient system $\dot{u}+\nabla_{\text {euc }} \mathscr{E}(u)=0$ in polar coordinates becomes

$$
\begin{aligned}
\dot{r}+\frac{\partial \tilde{\mathscr{E}}}{\partial r}(r, \theta) & =0 \\
\dot{\theta}+\frac{1}{r^{2}} \frac{\partial \tilde{\mathscr{E}}}{\partial \theta}(r, \theta) & =0
\end{aligned}
$$

Here, $\tilde{\mathscr{E}}(r, \theta)=\mathscr{E}(r \cos \theta, r \sin \theta)$. Let $r: \mathbb{R}_{+} \rightarrow(0, \infty)$ be a global solution of


Fig. 11.2 The level curves of the "mexican hat" of Palis \& de Melo near the unit circle

$$
\begin{equation*}
\dot{r}+\exp \left(\frac{-1}{r^{2}-1}\right) \frac{2 r\left(r^{2}-1\right)^{2}}{4 r^{4}+\left(r^{2}-1\right)^{4}}=0 \tag{11.4}
\end{equation*}
$$

with $r(0)>1$. Then $\lim _{t \rightarrow \infty} r(t)=1$ (exercise!). Moreover, if we put in addition

$$
\theta(t):=\frac{1}{r(t)^{2}-1},
$$

then a straightforward calculation shows that the pair $(r, \theta)$ is a solution of the gradient system above (in polar coordinates). However, for this solution one has

$$
\lim _{t \rightarrow \infty} r(t)=1 \quad \text { and } \quad \lim _{t \rightarrow \infty} \theta(t)=\infty
$$

Hence, the function $u: \mathbb{R}_{+} \rightarrow \mathbb{R}^{2}$ given by

$$
u(t)=(r(t) \cos \theta(t), r(t) \sin \theta(t)) \quad\left(t \in \mathbb{R}_{+}\right)
$$

is a global solution of the euclidean gradient system associated with $\mathscr{E}$, and has the whole unit circle as $\omega$-limit set. It "follows the gully", as described by Curry. Thus, in general, global and bounded solutions of gradient systems need not converge.

### 11.4 Exercises

11.1. Show that every global and bounded solution of a scalar ordinary differential equation converges to an equilibrium point.
11.2. Show that every maximal solution $r$ of (11.4) with $r(0)>1$ is global and satisfies $\lim _{t \rightarrow \infty} r(t)=1$
11.3 (Shooting method). Let $p>1$ and consider the set

$$
S:=\left\{\varphi \in H^{2}(0,1) \cap H_{0}^{1}(0,1):-\varphi^{\prime \prime}-|\varphi|^{p-1} \varphi=0\right\}
$$

of all equilibrium points of the problem

$$
\left\{\begin{array}{l}
u_{t}-u_{x x}-|u|^{p-1} u=0 \text { in }(0, \infty) \times(0,1), \\
u(t, 0)=u(t, 1)=0 \quad \text { for every } t \in(0, \infty) .
\end{array}\right.
$$

a) By recalling an existence and uniqueness result for ordinary differential equations, show that for every $c \in \mathbb{R}$ the initial-value problem

$$
\left\{\begin{array}{l}
-\varphi^{\prime \prime}-|\varphi|^{p-1} \varphi=0 \text { on } \mathbb{R},  \tag{11.5}\\
\varphi(0)=0 \\
\varphi^{\prime}(0)=c
\end{array}\right.
$$

admits a unique maximal solution $\varphi:\left[0, x_{\max }\right) \rightarrow \mathbb{R}$.
b) Show that the for every solution $\varphi$ of (11.5) the quantity $E\left(\varphi(x), \varphi^{\prime}(x)\right):=$ $\frac{1}{2} \varphi^{\prime}(x)^{2}+\frac{1}{p+1}|\varphi(x)|^{p+1}$ does not depend on $x$, that is,

$$
\begin{equation*}
\frac{1}{2} \varphi^{\prime}(x)^{2}+\frac{1}{p+1}|\varphi(x)|^{p+1}=\text { const }=: d \quad \text { for every } x \geq 0 \tag{11.6}
\end{equation*}
$$

Deduce from this that the maximal solution of (11.5) is global, that is, exists on $\mathbb{R}_{+}$. Moreover, express the constant $d$ in terms of $c$.
c) Let $\varphi$ be a solution of (11.5) and assume that $\varphi^{\prime}(0)=c>0$. Then, by continuity,

$$
x_{0}:=\inf \left\{x \geq 0: \varphi^{\prime}(x) \geq 0\right\}>0 .
$$

Express $x_{0}$ as a function of $c$.
Hint. On the interval $\left[0, x_{0}\right.$ ), one has $\varphi^{\prime}>0$. Equation (11.6) can thus be rewritten in the form

$$
\varphi^{\prime}(x)=\sqrt{2 d-\frac{2}{p+1}|\varphi(x)|^{p+1}} .
$$

Integrate this ordinary differential equation with separated variables.
d) Let $\varphi$ and $x_{0}$ be as in the previous item. Show that $\varphi\left(2 n x_{0}\right)=0$ for every integer $n \geq 1$. Show that $x_{0}=\frac{1}{2 n}$ for some integer $n \geq 1$ if and only if $\varphi$ (restricted to the interval $[0,1]$ ) is an equilibrium point.
e) Show that the mapping $S \rightarrow \mathbb{R}, \varphi \mapsto \varphi^{\prime}(0)$ is continuous and injective.
f) Show that the set $S$ is discrete.
11.4. Let $f: \mathbb{R} \rightarrow \mathbb{R}$ be a continuously differentiable function, and consider the semilinear diffusion equation

$$
\left\{\begin{array}{l}
u_{t}-u_{x x}+f(u)=0 \quad \text { in }(0, \infty) \times \mathbb{R} \\
\lim _{x \rightarrow-\infty} u(t, x)=\lim _{x \rightarrow+\infty} u(t, x)=0
\end{array}\right.
$$

Define $F(\zeta)=\int_{0}^{\zeta} f(s) d s$, let $\zeta_{0}:=\inf \{\zeta>0: F(\zeta) \leq 0\}$, and assume that

$$
\begin{aligned}
& f(0)=0 \text { and } f^{\prime}(0)>0 \\
& 0<\zeta_{0}<\infty \text { and } f\left(\zeta_{0}\right)<0
\end{aligned}
$$

a) Show that the problem

$$
\left\{\begin{array}{l}
-\varphi^{\prime \prime}+f(\varphi)=0 \quad \text { in } \mathbb{R} \\
\lim _{x \rightarrow-\infty} \varphi(x)=\lim _{x \rightarrow+\infty} \varphi(x)=0
\end{array}\right.
$$

admits a strictly positive solution (called ground state).
Hint. Similarly as in the previous exercise, one may employ a shooting method. Show that the initial value problem

$$
\left\{\begin{array}{l}
-\varphi^{\prime \prime}+f(\varphi)=0 \text { in } \mathbb{R} \\
\varphi(0)=\zeta \\
\varphi^{\prime}(0)=0
\end{array}\right.
$$

admits for appropriate $\zeta>0$ a global solution which is positive, symmetric with respect to the origin, and which satisfies $\lim _{x \rightarrow \pm \infty} \varphi(x)=0$. In order to study solutions of this initial value problem, it is convenient to note that the quantity $\frac{1}{2} \varphi^{\prime}(x)^{2}-F(\varphi(x))$ does not depend on $x$, and to study a scalar first order ordinary differential equation instead.
b) Show that the solution found in (a) belongs to $H^{1}(\mathbb{R})$.
c) Show that the semilinear diffusion equation above admits a continuum of equilibrium points in $H^{1}(\mathbb{R})$.
Hint. Consider translates of $\varphi$.

## Lecture 12

## Asymptotic behaviour of solutions of gradient systems II

In the years 1963 and 1965 appeared the following result of S. Łojasiewicz [Łojasiewicz (1963)], [Łojasiewicz (1965)].

Theorem 12.1 (Lojasiewicz). Let $\mathscr{E}: U \rightarrow \mathbb{R}$ be an analytic function defined on an open set $U \subseteq \mathbb{R}^{d}$, and let $\varphi \in U$ be an equilibrium point of $\mathscr{E}$. Then there exist constants $\theta \in\left(0, \frac{1}{2}\right], \sigma>0$ and $C \geq 0$ such that for every $u \in U$ with $\|u-\varphi\| \leq \sigma$ one has

$$
\begin{equation*}
|\mathscr{E}(u)-\mathscr{E}(\varphi)|^{1-\theta} \leq C\left\|\mathscr{E}^{\prime}(u)\right\|_{\left(\mathbb{R}^{d}\right)^{\prime}} \tag{12.1}
\end{equation*}
$$

The inequality (12.1), which may also be written in the form

$$
\left\|\left((\mathscr{E}(u)-\mathscr{E}(\varphi))^{\theta}\right)^{\prime}\right\|_{\left(\mathbb{R}^{d}\right)^{\prime}} \geq \frac{\theta}{C}>0
$$

is usually called Łojasiewicz inequality or, more precisely, Lojasiewicz gradient inequality. Theorem 12.1 expresses a particular, regular behaviour of analytic functions of several variables near equilibrium points, a behaviour which is in general not shared by $C^{\infty}$ functions. Its proof, however, is quite involved from the point of view of this lecture course, and we therefore omit it.

We start this lecture by stating the above theorem, because the Łojasiewicz inequality is a (or rather the?) fundamental tool in order to show convergence to equilibrium of global and bounded solutions of gradient systems. The discovery that the inequality may be applied in the context of gradient systems is due to Łojasiewicz himself. In his articles he proved that every global and bounded solution of a gradient system in $\mathbb{R}^{d}$ and with analytic energy converges to a single equilibrium. This result is in contrast to the mexican hat example of Palis and de Melo where the energy is $C^{\infty}(!)$ and the associated gradient system admits a global and bounded solution which does not converge.

We emphasize that one should really distinguish two independent problems in the context of the Łojasiewicz inequality and the convergence to equilibrium,
although Łojasiewicz' name is connected to both of them: the first problem is the problem of convergence to equilibrium of global solutions of gradient systems under the assumption that the Łojasiewicz inequality holds - this problem is basically solved with Theorem 12.2 below, be repeating essentially Łojasiewicz' idea. The second and more difficult problem is to know whether a concrete energy does satisfy the Łojasiewicz inequality near some particular equilibrium point. This problem is independent of the first one and independent of the formulation of a particular gradient system. A positive answer to the second question, that is, a proof of the Łojasiewicz inequality for a particular energy function, may have consequences for many different gradient systems, since out of one energy function one can generate a variety of gradient systems through the choice of the metric.

In this lecture we discuss both problems in two independent sections. In the first section we show that the infinite-dimensional variant of the Łojasiewicz inequality implies convergence to equilibrium of global and bounded solutions of abstract gradient systems. In the second section we prove the Łojasiewicz inequality in two comparatively simple situations.

### 12.1 The Łojasiewicz-Simon inequality and stabilisation of global solutions of gradient systems

Let $V$ be a Banach space, and let $H$ be a Hilbert space such that $V$ is densely and continuously embedded into $H$. Let $U \subseteq V$ be open, and let $\mathscr{E}: U \rightarrow \mathbb{R}$ be a continuously differentiable function. Let $g: U \rightarrow \operatorname{Inner}(H)$ be a metric on $U$. We assume that there exist constants $c_{1}, c_{2}>0$ such that, for every $v \in H$ and every $u \in U$,

$$
\begin{equation*}
c_{1}\|v\|_{H} \leq\|v\|_{g(u)} \leq c_{2}\|v\|_{H} \tag{12.2}
\end{equation*}
$$

Theorem 12.2. Let $u \in W_{\text {loc }}^{1,2}\left(\mathbb{R}_{+} ; H\right) \cap C\left(\mathbb{R}_{+} ; V\right)$ be a solution of the gradient system

$$
\begin{equation*}
\dot{u}+\nabla_{g} \mathscr{E}(u)=0 \tag{12.3}
\end{equation*}
$$

Assume that $u$ has relatively compact range in $U$, and assume that for every $t \in \mathbb{R}_{+}$ the energy equality

$$
\begin{equation*}
\int_{0}^{t}\|\dot{u}(s)\|_{g(u(s))}^{2} d s+\mathscr{E}(u(t))=\mathscr{E}(u(0)) \tag{12.4}
\end{equation*}
$$

holds. Assume further that there exists $\varphi \in \omega(u), \theta \in\left(0, \frac{1}{2}\right], \sigma>0$ and $C \geq 0$ such that for every $v \in U$ with $\|v-\varphi\|_{V} \leq \sigma$ one has

$$
\begin{equation*}
|\mathscr{E}(v)-\mathscr{E}(\varphi)|^{1-\theta} \leq C\left\|\mathscr{E}^{\prime}(v)\right\|_{V^{\prime}} \tag{12.5}
\end{equation*}
$$

Then $\lim _{t \rightarrow \infty}\|u(t)-\varphi\|_{V}=0$. Moreover, as $t \rightarrow \infty$,
12.1 The Łojasiewicz-Simon inequality and stabilisation of global solutions of gradient system31

$$
\|u(t)-\varphi\|_{H}= \begin{cases}O\left(e^{-c t}\right) & \text { for some } c>0, \text { if } \theta=\frac{1}{2}  \tag{12.6}\\ O\left(t^{-\theta /(1-2 \theta)}\right) & \text { if } \theta \in\left(0, \frac{1}{2}\right)\end{cases}
$$

In the following, the inequality (12.5) is called Łojasiewicz-Simon inequality, in order to honour also the work of L. Simon who generalized Łojasiewicz' ideas to energies and gradient systems on infinite-dimensional spaces. For example, Simon applied the Łojasiewicz-Simon inequality in order to prove convergence to equilibrium of solutions of semilinear diffusion equations with analytic energy [Simon (1983)].

Note that the energy equality (12.4) implies that the composite function $\mathscr{E}(u)$ belongs to $W_{\text {loc }}^{1,1}\left(\mathbb{R}_{+}\right)$, by Lemma 5.9. In the following, derivatives of the composite function are to be understood as weak derivatives. By (12.3), the energy equality may also be written in the form

$$
\begin{aligned}
& \int_{0}^{t}\left\|\nabla_{\mathscr{E}} \mathscr{E}(u(s))\right\|_{g(u(s))}^{2} d s+\mathscr{E}(u(t))=\mathscr{E}(u(0)) \quad \text { or } \\
& \int_{0}^{t}\left\|\nabla_{\mathscr{E}} \mathscr{E}(u(s))\right\|_{g(u(s))}\|\dot{u}(s)\|_{g(u(s))} d s+\mathscr{E}(u(t))=\mathscr{E}(u(0)) .
\end{aligned}
$$

Proof (of Theorem 12.2). By the energy equality (12.4), the function $\mathscr{E}$ is nonincreasing along $u$, and if $\mathscr{E}(u)$ is constant, then the function $u$ is constant. Since $\varphi \in \omega(u)$ by assumption, there exists an unbounded sequence $\left(t_{n}\right) \subseteq \mathbb{R}_{+}$such that $\lim _{n \rightarrow \infty} u\left(t_{n}\right)=\varphi$. Since $\mathscr{E}$ is continuous, we obtain $\lim _{n \rightarrow \infty} \mathscr{E}\left(u\left(t_{n}\right)\right)=\mathscr{E}(\varphi)$. Since $\mathscr{E}$ is nonincreasing along $u$, it follows that $\mathscr{E}(u(t)) \geq \mathscr{E}(\varphi)$ and $\lim _{t \rightarrow \infty} \mathscr{E}(u(t))=$ $\mathscr{E}(\varphi)$.

If $\mathscr{E}\left(u\left(t_{0}\right)\right)=\mathscr{E}(\varphi)$ for some $t_{0} \geq 0$, then $\mathscr{E}(u(t))=\mathscr{E}(\varphi)$ for every $t \geq t_{0}$, and therefore, by the energy equality (12.4), $\dot{u}(t)=0$ for $t \geq t_{0}$. In this case, the function $u$ is constant for $t \geq t_{0}$, and the assertions about convergence and decay estimate hold trivially.

So we can assume that $\mathscr{E}(u(t))>\mathscr{E}(\varphi)$ for every $t \geq 0$. Let, for every $t \geq 0$,

$$
\mathscr{H}(t):=(\mathscr{E}(u(t))-\mathscr{E}(\varphi))^{\theta} .
$$

Then $\mathscr{H}$ is nonincreasing, $\mathscr{H}(t)>0$ for every $t \geq 0$, and $\lim _{t \rightarrow \infty} \mathscr{H}(t)=0$. Let $t_{0} \geq 0$ be such that $\left\|u\left(t_{0}\right)-\varphi\right\|_{V}<\sigma$, and define

$$
t_{1}:=\inf \left\{t \geq t_{0}:\|u(t)-\varphi\|_{V}=\sigma\right\} .
$$

By continuity of the function $u$, we have $t_{1}>t_{0}$. By using successively the chain rule, the energy equality (12.4) and the fact that $u$ is a solution of (12.3), and the assumption (12.2), we obtain for almost every $t \in\left[t_{0}, t_{1}\right)$,

$$
\begin{aligned}
-\frac{d}{d t} \mathscr{H}(t) & =\theta(\mathscr{E}(u(t))-\mathscr{E}(\varphi))^{\theta-1}\left(-\frac{d}{d t} \mathscr{E}(u(t))\right) \\
& =\theta(\mathscr{E}(u(t))-\mathscr{E}(\varphi))^{\theta-1}\left\|\nabla_{g} \mathscr{E}(u(t))\right\|_{g(u(t))}\|\dot{u}(t)\|_{g(u(t))} \\
& \geq \theta c_{1}(\mathscr{E}(u(t))-\mathscr{E}(\varphi))^{\theta-1}\left\|\nabla_{g} \mathscr{E}(u(t))\right\|_{g(u(t))}\|\dot{u}(t)\|_{H}
\end{aligned}
$$

By using the assumption (12.2) again, we obtain that for every $t \in \mathbb{R}_{+}$the estimate

$$
\begin{aligned}
\left\|\nabla_{g} \mathscr{E}(u(t))\right\|_{g(u(t))} & =\sup _{\|v\|_{g(u(t))} \leq 1}\left\langle\nabla_{g} \mathscr{E}(u(t)), v\right\rangle_{g(u(t))} \\
& =\sup _{\|v\|_{g(u(t))} \leq 1}\left\langle\mathscr{E}^{\prime}(u(t)), v\right\rangle_{V^{\prime}, V} \\
& \geq \sup _{c_{2}\|\nu\|_{H} \leq 1}\left\langle\mathscr{E}^{\prime}(u(t)), v\right\rangle_{V^{\prime}, V} \\
& \geq \sup _{c_{3} c_{2}\|v\|_{V} \leq 1}\left\langle\mathscr{E}^{\prime}(u(t)), v\right\rangle_{V^{\prime}, V} \\
& =\frac{1}{c_{3} c_{2}}\left\|\mathscr{E}^{\prime}(u(t))\right\|_{V^{\prime}}
\end{aligned}
$$

where $c_{3}>0$ is the constant of the embedding $V \hookrightarrow H$. This estimate and the Łojasiewicz-Simon inequality (12.5) imply that, for almost every $t \in\left[t_{0}, t_{1}\right.$ ),

$$
\begin{align*}
-\frac{d}{d t} \mathscr{H}(t) & \geq \frac{\theta c_{1}}{c_{2} c_{3}}(\mathscr{E}(u(t))-\mathscr{E}(\varphi))^{\theta-1}\left\|\mathscr{E}^{\prime}(u(t))\right\|_{V^{\prime}}\|\dot{u}(t)\|_{H}  \tag{12.7}\\
& \geq \frac{\theta c_{1}}{C c_{2} c_{3}}\|\dot{u}(t)\|_{H}
\end{align*}
$$

Hence, for almost every $t \in\left[t_{0}, t_{1}\right)$,

$$
\begin{align*}
\|u(t)-\varphi\|_{H} & \leq\left\|u(t)-u\left(t_{0}\right)\right\|_{H}+\left\|u\left(t_{0}\right)-\varphi\right\|_{H}  \tag{12.8}\\
& \leq \int_{t_{0}}^{t}\|\dot{u}(s)\|_{H} d s+\left\|u\left(t_{0}\right)-\varphi\right\|_{H} \\
& \leq \frac{C c_{2} c_{3}}{\theta c_{1}} \mathscr{H}\left(t_{0}\right)+\left\|u\left(t_{0}\right)-\varphi\right\|_{H} .
\end{align*}
$$

Now let $\left(t_{0}^{n}\right) \subseteq \mathbb{R}_{+}$be an unbounded, increasing sequence such that $\sigma>\| u\left(t_{0}^{n}\right)-$ $\varphi \|_{V} \rightarrow 0$, and define the corresponding $t_{1}^{n}$ as above. Assume that $t_{1}^{n}$ was finite for every $n$. Then, by definition of $t_{1}^{n}$ and by continuity of $u,\left\|u\left(t_{1}^{n}\right)-\varphi\right\|_{V}=\sigma$ for every $n$. Since $u$ has relatively compact range in $U \subseteq V$, we can extract a subsequence of $\left(t_{1}^{n}\right)$ (which we denote for simplicity again by $\left(t_{1}^{n}\right)$ ) such that $\lim _{n \rightarrow \infty} u\left(t_{1}^{n}\right)=: \psi$. By continuity of the norm, $\|\psi-\varphi\|_{V}=\sigma>0$. On the other hand, from the inequality (12.8) we obtain $\|\psi-\varphi\|_{H}=\lim _{n \rightarrow \infty}\left\|u\left(t_{1}^{n}\right)-\varphi\right\|_{H}=0$, which is a contradiction.

Hence, for some $n$ large enough, $t_{1}^{n}=+\infty$. By (12.7), this implies $\dot{u} \in L^{1}\left(\mathbb{R}_{+} ; H\right)$. By Cauchy's criterion, $\lim _{t \rightarrow \infty} u(t)$ exists in $H$. By using the relative compactness of the range of $u$ in $V$, a subsubsequence argument, and since $\varphi \in \omega(u)$, this implies $\lim _{t \rightarrow \infty} u(t)=\varphi$ in $V$.

In order to prove the decay estimate, let $t_{0} \geq 0$ be large enough such that $\| u(t)-$ $\varphi \|_{V} \leq \sigma$ for every $t \geq t_{0}$. By the inequality (12.7), for every $t \geq t_{0}$,

$$
\begin{align*}
\|u(t)-\varphi\|_{H} & \leq \int_{t}^{\infty}\|\dot{u}(s)\|_{H} d s  \tag{12.9}\\
& \leq \frac{C c_{2} c_{3}}{\theta c_{1}} \mathscr{H}(t)
\end{align*}
$$

Moreover, by the energy equality (12.4), for almost every $t \geq t_{0}$,

$$
\begin{aligned}
-\frac{d}{d t} \mathscr{H}(t) & =\theta(\mathscr{E}(u(t))-\mathscr{E}(\varphi))^{\theta-1}\left(-\frac{d}{d t} \mathscr{E}(u(t))\right) \\
& =\theta(\mathscr{E}(u(t))-\mathscr{E}(\varphi))^{\theta-1}\left\|\nabla_{g} \mathscr{E}(u(t))\right\|_{g(u(t))}^{2} \\
& =\frac{\theta}{c_{2}^{2} c_{3}^{2}}(\mathscr{E}(u(t))-\mathscr{E}(\varphi))^{\theta-1}\left\|\mathscr{E}^{\prime}(u(t))\right\|_{V}^{2} \\
& \geq \frac{\theta}{C^{2} c_{2}^{2} c_{3}^{2}}(\mathscr{E}(u(t))-\mathscr{E}(\varphi))^{1-\theta} \\
& =\frac{\theta}{C^{2} c_{2}^{2} c_{3}^{2}} \mathscr{H}(t)^{\frac{1-\theta}{\theta}}
\end{aligned}
$$

Now, if we integrate the resulting inequality

$$
-\frac{d}{d t} \mathscr{H}(t) \mathscr{H}(t)^{-\frac{1-\theta}{\theta}}=\left\{\begin{array}{ll}
-\frac{d}{d t}(\log \mathscr{H}(t)) & \text { if } \theta=\frac{1}{2} \\
\frac{\theta}{1-2 \theta} \frac{d}{d t} \mathscr{H}(t)^{-\frac{1-2 \theta}{\theta}} & \text { if } \theta \in\left(0, \frac{1}{2}\right)
\end{array}\right\} \geq \frac{\theta}{C^{2} c_{2}^{2} c_{3}^{2}}
$$

over the interval $\left(t_{0}, t\right)$, then we obtain the estimate

$$
\mathscr{H}(t)= \begin{cases}O\left(e^{-c t}\right) & \text { if } \theta=\frac{1}{2} \\ O\left(t^{-\theta /(1-2 \theta)}\right) & \text { if } \theta \in\left(0, \frac{1}{2}\right)\end{cases}
$$

Combining this estimate with (12.9), the decay estimate for $\|u-\varphi\|_{H}$ follows.

### 12.2 The Łojasiewicz-Simon inequality in Hilbert spaces

Let $V$ be a Hilbert space and let $U \subseteq V$ be an open subset. Let $\mathscr{E} \in C^{2}(U)$, and let $\varphi \in U$ be an equilibrium point of $\mathscr{E}$, that is, $\mathscr{E}^{\prime}(\varphi)=0$. We formulate conditions which imply that $\mathscr{E}$ satisfies the Łojasiewicz-Simon inequality near $\varphi$, that is, that there exists $\theta \in\left(0, \frac{1}{2}\right], \sigma>0$ and $C \geq 0$ such that for every $u \in U$ with $\|u-\varphi\|_{V} \leq \sigma$ one has

$$
\begin{equation*}
|\mathscr{E}(u)-\mathscr{E}(\varphi)|^{1-\theta} \leq C\left\|\mathscr{E}^{\prime}(u)\right\|_{V^{\prime}} \tag{12.10}
\end{equation*}
$$

Although in this inequality only the first derivative of $\mathscr{E}$ appears, the conditions and arguments in this section involve the second derivative, too. Recall that the Fréchet derivative $\mathscr{E}^{\prime}$ maps $V$ into $V^{\prime}$, and that the second derivative at a point $u \in U$ is thus a linear operator from $V$ into $V^{\prime}$.

Theorem 12.3. Assume that $\mathscr{E}^{\prime \prime}(\varphi)$ is continuously invertible. Then $\mathscr{E}$ satisfies the Łojasiewicz-Simon inequality near $\varphi$ with $\theta=\frac{1}{2}$.

Proof. Consider the Taylor expansions of $\mathscr{E}$ and $\mathscr{E}^{\prime}$,

$$
\begin{aligned}
\mathscr{E}(u) & =\mathscr{E}(\varphi)+\mathscr{E}^{\prime}(\varphi)(u-\varphi)+O\left(\|u-\varphi\|_{V}^{2}\right) \quad \text { and } \\
\mathscr{E}^{\prime}(u) & =\mathscr{E}^{\prime}(\varphi)+\mathscr{E}^{\prime \prime}(\varphi)(u-\varphi)+o\left(\|u-\varphi\|_{V}\right)
\end{aligned}
$$

Since $\mathscr{E}^{\prime}(\varphi)=0$, the first one yields

$$
|\mathscr{E}(u)-\mathscr{E}(\varphi)| \leq C\|u-\varphi\|_{V}^{2}
$$

in a neighbourhood of $\varphi$. Since $\mathscr{E}^{\prime \prime}(\varphi)$ is invertible, one has $\left\|\mathscr{E}^{\prime \prime}(\varphi)(u-\varphi)\right\|_{V^{\prime}} \geq$ $c\|u-\varphi\|_{V}$ for every $u \in U$ and some constant $c>0$. Hence, the second Taylor expansion yields

$$
\left\|\mathscr{E}^{\prime}(u)\right\|_{V^{\prime}} \geq \frac{c}{2}\|u-\varphi\|_{V}
$$

in a neighbourhood of $\varphi$. Combining the preceding two inequalities yields

$$
|\mathscr{E}(u)-\mathscr{E}(\varphi)|^{\frac{1}{2}} \leq C\left\|\mathscr{E}^{\prime}(u)\right\|_{V^{\prime}}
$$

in a neighbourhood of $\varphi$, and this is the claim.
The short proof of Theorem 12.3 relies only on Taylor expansion and the definition of the Fréchet derivative. This idea can be generalized to situations in which $\mathscr{E}^{\prime \prime}(\varphi)$ is not necessarily invertible, but in which, by means of the implicit function theorem and a nonlinear decomposition of the space $V$, one can reduce the arguments to the invertible case. In the rest of this section, we assume that $\mathscr{E}^{\prime \prime}(\varphi)$ is a Fredholm operator, that is, the kernel $\operatorname{ker} \mathscr{E}^{\prime \prime}(\varphi)=\left\{u \in V: \mathscr{E}^{\prime \prime}(\varphi) u=0\right\}$ is finite dimensional and the range $\operatorname{rg} \mathscr{E}^{\prime \prime}(\varphi)=\left\{\mathscr{E}^{\prime \prime}(\varphi) u: u \in V\right\}$ is closed in $V^{\prime}$ and has finite codimension.

Let $P \in \mathscr{L}(V)$ be any projection onto $\operatorname{ker} \mathscr{E}^{\prime \prime}(\varphi)$. Then $V$ is the direct topological sum

$$
\begin{aligned}
V & =V_{0} \oplus V_{1} \\
& =\operatorname{rg} P \oplus \operatorname{ker} P \\
& =\operatorname{ker} \mathscr{E}^{\prime \prime}(\varphi) \oplus \operatorname{ker} P
\end{aligned}
$$

Moreover, the spaces $\operatorname{rg} P^{\prime}, \operatorname{ker} P^{\prime} \subseteq V^{\prime}$, where $P^{\prime} \in \mathscr{L}\left(V^{\prime}\right)$ is the adjoint projection, may be naturally identified with the dual spaces $V_{0}^{\prime}$ and $V_{1}^{\prime}$, respectively, and we
simply write $V_{1}^{\prime}=\operatorname{ker} P^{\prime}$.
Symmetry of the second derivative of real-valued functions is a well known property of twice continuously differentiable functions defined on $\mathbb{R}^{d}$; see [Cartan (1967), Théorème 5.1.1] for the analogous result for functions defined on a Banach space.

Theorem 12.4 (Schwarz). For every $u \in U$ the linear operator $\mathscr{E}^{\prime \prime}(u): V \rightarrow V^{\prime}$ is symmetric, that is, for every $v, w \in V$ one has

$$
\left\langle\mathscr{E}^{\prime \prime}(u) v, w\right\rangle_{V^{\prime}, V}=\left\langle\mathscr{E}^{\prime \prime}(u) w, v\right\rangle_{V^{\prime}, V}
$$

By Schwarz' theorem, for every $u, v \in V$,

$$
\begin{aligned}
\left\langle P^{\prime} \mathscr{E}^{\prime \prime}(\varphi) u, v\right\rangle_{V^{\prime}, V} & =\left\langle\mathscr{E}^{\prime \prime}(\varphi) u, P v\right\rangle_{V^{\prime}, V} \\
& =\left\langle\mathscr{E}^{\prime \prime}(\varphi) P v, u\right\rangle_{V^{\prime}, V} \\
& =0
\end{aligned}
$$

where in the last inequality we have used that $P$ projects onto the kernel of $\mathscr{E}^{\prime \prime}(\varphi)$. This equality implies that

$$
\begin{equation*}
\operatorname{rg} \mathscr{E}^{\prime \prime}(\varphi) \subseteq \operatorname{ker} P^{\prime}=V_{1}^{\prime} \tag{12.11}
\end{equation*}
$$

Lemma 12.5. The linear operator $\mathscr{E}^{\prime \prime}(\varphi): V_{1} \rightarrow V_{1}^{\prime}$ is continuously invertible.
Proof. Since $V_{1} \cap V_{0}=V_{1} \cap \operatorname{ker} \mathscr{E}^{\prime \prime}(\varphi)=\{0\}$, the operator $\mathscr{E}^{\prime \prime}(\varphi)$ is injective on $V_{1}$.
We next prove that $\mathscr{E}^{\prime \prime}(\varphi)$ has dense range in $V_{1}^{\prime}$. Note that $\mathscr{E}^{\prime \prime}(\varphi)\left(V_{1}\right)=$ $\mathscr{E}^{\prime \prime}(\varphi)(V)$. Hence, in order to show that $\mathscr{E}^{\prime \prime}(\varphi)$ has dense range in $V_{1}^{\prime}$, it suffices show that

$$
\left.\begin{array}{l}
u \in V_{1} \text { and } \\
\left\langle\mathscr{E}^{\prime \prime}(\varphi) v, u\right\rangle_{V^{\prime}, V}=0 \text { for every } v \in V
\end{array}\right\} \Rightarrow u=0
$$

This implication, however, follows immediately from Schwarz' theorem and injectivity of $\mathscr{E}^{\prime \prime}(\varphi)$ on $V_{1}$.

By assumption, $\mathscr{E}^{\prime \prime}(\varphi)$ has closed range in $V^{\prime}$, so that $\mathscr{E}^{\prime \prime}(\varphi)$ is surjective, hence bijective from $V_{1}$ onto $V_{1}^{\prime}$. The assertion of the lemma follows now from the bounded inverse theorem.

Lemma 12.6. Let $P \in \mathscr{L}(V)$ be a projection onto $\operatorname{ker} \mathscr{E}^{\prime \prime}(\varphi)$, and define the set

$$
S:=\left\{u \in U:\left(I-P^{\prime}\right) \mathscr{E}^{\prime}(u)=0\right\}
$$

Then, locally near $\varphi$, S a differentiable manifold, called critical manifold, satisfying

$$
\operatorname{dim} S=\operatorname{dim} \operatorname{ker} \mathscr{E}^{\prime \prime}(\varphi)
$$

If $\mathscr{E} \in C^{k}(U)$ for some $k \geq 2$, then $S$ is a $C^{k-1}$-manifold. If $\mathscr{E}$ is analytic, then $S$ is analytic.

Proof. Consider the function

$$
\begin{aligned}
G: V= & V_{0} \oplus V_{1} \supseteq U \\
& \rightarrow V_{1}^{\prime}, \\
& u=u_{0}+u_{1} \mapsto\left(I-P^{\prime}\right) \mathscr{E}^{\prime}(u) .
\end{aligned}
$$

Since $\mathscr{E}^{\prime}$ is continuously differentiable, the function $G$ is continuously differentiable, too. Moreover, $G(\varphi)=G\left(\varphi_{0}+\varphi_{1}\right)=0$ and $G^{\prime}(\varphi)=\left(I-P^{\prime}\right) \mathscr{E}^{\prime \prime}(\varphi)=\mathscr{E}^{\prime \prime}(\varphi)$ (see (12.11) for the last equality). By Lemma 12.5, the partial derivative $\frac{\partial G}{\partial u_{1}}(\varphi)=$ $\left.\mathscr{E}^{\prime \prime}(\varphi)\right|_{V_{1}}: V_{1} \rightarrow V_{1}^{\prime}$ is an isomorphism (continuously invertible). Hence, by the implicit function theorem (Theorem 9.2), there exists a neighbourhood $U_{0} \subseteq V_{0}$ of $\varphi_{0}$, a neighbourhood $U_{1} \subseteq V_{1}$ of $\varphi_{1}, U_{0}+U_{1} \subseteq U$, and a function $g \in C^{1}\left(U_{0} ; U_{1}\right)$ such that $g\left(\varphi_{0}\right)=\varphi_{1}$ and

$$
\left\{u \in U_{0}+U_{1}: G(u)=0\right\}=\left\{\left(u_{0}, g\left(u_{0}\right)\right): u_{0} \in U_{0}\right\}
$$

By definition of the function $G$ and the set $S$, the set on the left-hand side of this equality is just the intersection of $S$ with the neighbourhood $U_{0}+U_{1}$ of $\varphi$. So, locally near $\varphi, S$ is the graph of the differentiable function $g$, that is, $S$ is a differentiable manifold. Higher regularity of the manifold $S$ in the case of higher regularity of $\mathscr{E}$ follows immediately from the implicit function theorem. The lemma is proved.

Note that the critical manifold may not be a submanifold of $V$, but by Lemma 12.6 it is locally near $\varphi$ a submanifold (it is the graph of the implicit function $g$ ). The critical manifold depends on the choice of the projection $P$, but it always contains the set of all equilibrium points $S_{0}$ :

$$
S_{0}:=\left\{u \in U: \mathscr{E}^{\prime}(u)=0\right\} \subseteq S
$$

Theorem 12.7. Define the critical manifold as in Lemma 12.6, and assume that the restriction $\left.\mathscr{E}\right|_{S}$ satisfies the Łojasiewicz-Simon inequality near $\varphi$, that is, there exist constants $\theta \in\left(0, \frac{1}{2}\right], \sigma>0$ and $C \geq 0$ such that for every $u \in U \cap S$ (!) with $\|u-\varphi\|_{V} \leq \sigma$ one has

$$
|\mathscr{E}(u)-\mathscr{E}(\varphi)|^{1-\theta} \leq C\left\|\mathscr{E}^{\prime}(u)\right\|_{V^{\prime}}
$$

Then $\mathscr{E}$ itself satisfies the Łojasiewicz-Simon inequality near $\varphi$ with the same exponent $\theta$.

Proof. Choose the neighbourhood $U:=U_{0}+U_{1}$ of $\varphi$ and the implicit function $g$ : $U_{0} \rightarrow U_{1}$ as in the proof of Lemma 12.6. Suppose that $U$ is sufficiently small so that the restriction $\left.\mathscr{E}\right|_{S}$ satisfies the Łojasiewicz-Simon inequality in $U \cap S$.

We define a nonlinear projection $Q: U \rightarrow U$ by

$$
Q u=Q\left(u_{0}+u_{1}\right):=u_{0}+g\left(u_{0}\right) .
$$

Note that, for every $u \in U$,

$$
\begin{aligned}
& Q u \in S \text { and } \\
& u-Q u \in V_{1}
\end{aligned}
$$

Moreover, $Q \varphi=\varphi$. For every $u \in U$, the Taylor expansion of $\mathscr{E}$ at $Q u$ is

$$
\begin{aligned}
& \mathscr{E}(u)-\mathscr{E}(Q u)= \\
& \quad=\left\langle\mathscr{E}^{\prime}(Q u), u-Q u\right\rangle_{V^{\prime}, V}+\frac{1}{2}\left\langle\mathscr{E}^{\prime \prime}(Q u)(u-Q u), u-Q u\right\rangle_{V^{\prime}, V}+o\left(\|u-Q u\|^{2}\right) .
\end{aligned}
$$

By definition of $V_{1}$, and by definition of the manifold $S$,

$$
\begin{aligned}
\left\langle\mathscr{E}^{\prime}(Q u), u-Q u\right\rangle_{V^{\prime}, V} & =\left\langle\mathscr{E}^{\prime}(Q u),(I-P)(u-Q u)\right\rangle_{V^{\prime}, V} \\
& =\left\langle\left(I-P^{\prime}\right) \mathscr{E}^{\prime}(Q u), u-Q u\right\rangle_{V^{\prime}, V} \\
& =0
\end{aligned}
$$

that is, the first term on the right-hand side of the Taylor expansion of $\mathscr{E}$ is zero. Therefore, and since $\mathscr{E}^{\prime \prime}$ is uniformly bounded on $U$ by continuity, if we choose $U$ small enough, we have for every $u \in U$

$$
\begin{equation*}
|\mathscr{E}(u)-\mathscr{E}(Q u)| \leq C\|u-Q u\|_{V}^{2} \tag{12.12}
\end{equation*}
$$

From now on, the constant $C$ may vary from line to line. By the definition of differentiability,

$$
\begin{equation*}
\mathscr{E}^{\prime}(u)-\mathscr{E}^{\prime}(Q u)=\mathscr{E}^{\prime \prime}(Q u)(u-Q u)+o(\|u-Q u\|) \tag{12.13}
\end{equation*}
$$

We apply the projection $I-P^{\prime}$ to this equality and use that $Q u \in S$ in order to obtain

$$
\begin{equation*}
\left(I-P^{\prime}\right) \mathscr{E}^{\prime}(u)=\left(I-P^{\prime}\right) \mathscr{E}^{\prime \prime}(Q u)(u-Q u)+o(\|u-Q u\|) \tag{12.14}
\end{equation*}
$$

By Lemma 12.5, the operator $\left(I-P^{\prime}\right) \mathscr{E}^{\prime \prime}(\varphi)=\mathscr{E}^{\prime \prime}(\varphi): V_{1} \rightarrow V_{1}^{\prime}$ is continuously invertible. Hence, by continuity and if we choose $U$ small enough, then $\left(I-P^{\prime}\right) \mathscr{E}^{\prime \prime}(Q u): V_{1} \rightarrow V_{1}^{\prime}$ is continuously invertible for all $u \in U$ and the inverses are uniformly bounded in $U$. Hence, by (12.14), there exists a constant $C \geq 0$ such that for every $u \in U$

$$
\begin{equation*}
\|u-Q u\|_{V} \leq C\left\|\left(I-P^{\prime}\right) \mathscr{E}^{\prime}(u)\right\|_{V^{\prime}} \leq C\left\|\mathscr{E}^{\prime}(u)\right\|_{V^{\prime}} \tag{12.15}
\end{equation*}
$$

By (12.13) and (12.15),

$$
\begin{equation*}
\left\|\mathscr{E}^{\prime}(Q u)\right\|_{V^{\prime}} \leq\left\|\mathscr{E}^{\prime}(u)\right\|_{V^{\prime}}+C\|u-Q u\|_{V} \leq C\left\|\mathscr{E}^{\prime}(u)\right\|_{V^{\prime}} \tag{12.16}
\end{equation*}
$$

Combining the estimates (12.12) and (12.15) with the assumption that $\left.\mathscr{E}\right|_{S}$ satisfies the Łojasiewicz-Simon inequality in $U \cap S$, and using also the estimate (12.16), we obtain that for every $u \in U$

$$
\begin{aligned}
|\mathscr{E}(u)-\mathscr{E}(\varphi)| & \leq|\mathscr{E}(u)-\mathscr{E}(Q u)|+|\mathscr{E}(Q u)-\mathscr{E}(\varphi)| \\
& \leq C\left\|\mathscr{E}^{\prime}(u)\right\|_{V^{\prime}}^{2}+C\left\|\mathscr{E}^{\prime}(Q u)\right\|_{V^{\prime}}^{1 /(1-\theta)} \\
& \leq C\left(\left\|\mathscr{E}^{\prime}(u)\right\|_{V^{\prime}}^{2}+\left\|\mathscr{E}^{\prime}(u)\right\|_{V^{\prime}}^{1 /(1-\theta)}\right.
\end{aligned}
$$

Choosing $U$ sufficiently small, we can by continuity assume that $\left\|\mathscr{E}^{\prime}(u)\right\|_{V^{\prime}} \leq 1$ for every $u \in U$. Since $\theta \in\left(0, \frac{1}{2}\right]$, we then obtain

$$
|\mathscr{E}(u)-\mathscr{E}(\varphi)| \leq C\left\|\mathscr{E}^{\prime}(u)\right\|_{V^{\prime}}^{1 /(1-\theta)} \text { for every } u \in U
$$

and this is the claim of Theorem 12.7.
Corollary 12.8. Let $\mathscr{E} \in C^{2}(U)$, let $\varphi \in V$ be an equilibrium point, and assume that $\mathscr{E}^{\prime \prime}(\varphi)$ is a Fredholm operator. Define the critical manifold $S$ as in Lemma 12.6. Assume that the set of all equilibrium points,

$$
S_{0}:=\left\{u \in V: \mathscr{E}^{\prime}(u)=0\right\}
$$

forms a neighbourhood of $\varphi$ in $S$. Then $\mathscr{E}$ satisfies the Łojasiewicz-Simon inequality with exponent $\theta=\frac{1}{2}$.

Proof. By assumption, the derivative $\mathscr{E}$ © is zero in a neighbourhood of $\varphi$ in $S$. This implies that the restriction $\left.\mathscr{E}\right|_{S}$ is constant in the same neighbourhood. A constant function trivially satisfies the Łojasiewicz-Simon inequality for the Łojasiewicz exponent $\theta=\frac{1}{2}$. The claim follows from Theorem 12.7.

### 12.3 Exercises

12.1. a) Let $F: \mathbb{R} \rightarrow \mathbb{R}$ be a $k$ times continuously differentiable function $(k \geq 2)$ such that $F(0)=F^{\prime}(0)=\cdots=F^{(k-1)}(0)=0$ and $F^{(k)}(0) \neq 0$. Show that $F$ satisfies the Łojasiewicz inequality near 0 with exponent $\theta=\frac{1}{k}$. Deduce that every analytic function $F: \mathbb{R} \rightarrow \mathbb{R}$ satisfies the Łojasiewicz inequality near every equilibrium point.
b) Let $F(u)=u^{k}$ for some $k \geq 2$. Solve the differential equation $\dot{u}+F^{\prime}(u)=$ 0 , show that every solution satisfies $\lim _{t \rightarrow \infty} u(t)=0$, and compare the actual decay rate of $u$ with the decay rate obtained by Theorem 12.2 and (a).
Remark. The proofs in (a) rely on Taylor expansions and should be elementary. However, the proofs in the one-dimensional case considered here do not give a real hint how to prove the corresponding result in higher dimensions like, for example, Theorem 12.1.
12.2. a) Let $V$ be a Hilbert space and $\mathscr{E}: V \rightarrow \mathbb{R}$ be a continuous, coercive, quadratic form. Show that $\mathscr{E}^{\prime \prime}(0)$ is invertible.
Hint. Use the Riesz-Fréchet theorem (Theorem D.53).
b) Let $c: \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function such that $\lim _{x \rightarrow \pm \infty} c(x)=0$. Show that the (linear) multiplication operator

$$
\begin{aligned}
H^{1}(\mathbb{R}) & \rightarrow L^{2}(\mathbb{R}), \\
u & \mapsto c u
\end{aligned}
$$

is compact.
Hint. One may use that for every bounded interval $(a, b)$, the embedding $H^{1}(a, b) \hookrightarrow L^{2}(a, b)$ is compact (see Exercise 7.1).
c) Let $c: \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function such that $\lim _{x \rightarrow \pm \infty} c(x)=c_{0}>0$, and consider the quadratic form

$$
\begin{aligned}
\mathscr{E}: H^{1}(\mathbb{R}) & \rightarrow \mathbb{R} \\
u & \mapsto \frac{1}{2} \int_{\mathbb{R}}\left(u_{x}^{2}+c u^{2}\right) .
\end{aligned}
$$

Show that $\mathscr{E}^{\prime \prime}(0)$ is a Fredholm operator.
Hint. By the Riesz-Schauder Theorem, every operator of the form $I+K$ with $K$ compact, is a Fredholm operator.
12.3. We recall the setting of Exercise 11.4. Let $f: \mathbb{R} \rightarrow \mathbb{R}$ be a continuously differentiable function, define $F(\zeta):=\int_{0}^{\zeta} f(s) d s$, let $\zeta_{0}:=\inf \{\zeta>0: F(\zeta) \leq 0\}$, and assume that

$$
\begin{aligned}
& f(0)=0 \text { and } f^{\prime}(0)>0 \\
& 0<\zeta_{0}<\infty \text { and } f\left(\zeta_{0}\right)<0
\end{aligned}
$$

Consider the energy

$$
\begin{aligned}
\mathscr{E}: H^{1}(\mathbb{R}) & \rightarrow \mathbb{R}, \\
u & \mapsto \frac{1}{2} \int_{\mathbb{R}} u_{x}^{2}+\int_{\mathbb{R}} F(u) .
\end{aligned}
$$

a) Show that $\mathscr{E}^{\prime \prime}(0)$ is invertible. Conclude that $\mathscr{E}$ satisfies the ŁojasiewiczSimon inequality near 0 with exponent $\theta=\frac{1}{2}$.
b) Let $\varphi$ be the ground state found in Exercise 11.4 (a), that is, a solution $\varphi \geq 0$, $\varphi \neq 0$ of the problem

$$
\left\{\begin{array}{l}
-\varphi^{\prime \prime}+f(\varphi)=0 \quad \text { in } \mathbb{R} \\
\lim _{x \rightarrow \pm \infty} \varphi(x)=0
\end{array}\right.
$$

Show that $\mathscr{E}^{\prime \prime}(\varphi)$ is a Fredholm operator.
c) In addition to (b), show that the kernel of $\mathscr{E}^{\prime \prime}(\varphi)$ is one-dimensional.

Hint. Show that $\psi \in C^{2}(\mathbb{R}) \cap H^{1}(\mathbb{R})$ belongs to the kernel of $\mathscr{E}^{\prime \prime}(\varphi)$ if and
only if

$$
\left\{\begin{array}{l}
-\psi^{\prime \prime}+f^{\prime}(\varphi) \psi=0 \text { in } \mathbb{R} \\
\lim _{x \rightarrow \pm \infty} \psi(x)=0
\end{array}\right.
$$

One solution is $\psi=\varphi^{\prime}$. A second solution $\psi$ may be found by substitution $z:=\left(\frac{\psi}{\varphi^{\prime}}\right)^{\prime}$ which leads to a first order differential equation.
d) Show that, locally near $\varphi$, the set of all equilibrium points $S_{0}=\left\{u \in H^{1}(\mathbb{R})\right.$ : $\left.\mathscr{E}^{\prime}(u)=0\right\}$ and the critical manifold $S$ defined in Lemma 12.6 coincide. Conclude that $\mathscr{E}$ satisfies the Łojasiewicz-Simon inequality near 0 with exponent $\theta=\frac{1}{2}$.

## Lecture 13

The universe, soap bubbles and curve shortening

I do not suppose that there is any one in this room who has not occasionally blown a common soap-bubble, and while admiring the perfection of its form, and the marvellous brilliancy of its colours, wondered how it is that such a magnificent object can be so easily produced. I hope that none of you are yet tired of playing with bubbles, ...
C. V. Boys, author of Soap-bubbles and the Forces which Mould Them, 1896

My grandfather talked continuously about soap bubbles, and of course in mathematical terms. I did not understand a word of what he said.

Bernhard Caesar Einstein, the grandson of Albert Einstein
To be clear from the very beginning of this last ISEM lecture: there will be no mathematical discussion of any models describing the evolution of the universe or of soap bubbles in the air. However, the title shall suggest that there is an important class of evolution problems in physics, chemistry or engineering dealing with moving curves and surfaces. Some of these models are gradient systems.


Soap bubbles, Jean Siméon Chardin, mid-18th century

Moving curves and surfaces can be observed in everyday's life, for example in the form of soap bubbles blown into the air and moving away by wind or gravitational forces. However, these soap bubbles are not only moving as a whole through the air, but they exhibit also inner deformations due to various surface forces acting in the soap film. Especially for large bubbles such deformations can be observed with the eye; smaller bubbles seem to exist only as round spheres. Everybody who is not yet tired of playing with bubbles, can repeat this experiment, and everybody who is curious to understand this model in a deeper way may formulate the evolution equations behind it and analyze it from a mathematical point of view. Besides soap bubbles, vibrating strings or plates, the water surface on a lake or in a glass of water, the sharp interface separating two chemical
phases, or the interface between a piece of ice and surrounding water are further examples of surfaces which change their shape with time and which may be easily observed in our daily life. The description of the physical and chemical forces or processes acting on / in these surfaces leads to partial differential equations, and the mathematical analysis of these partial differential equations may allow us to obtain a deeper understanding of the various models.

In the theory of relativity or gravitational physics, one deals with a surface at a much different scale at which our human intuition is of no big help. If we believe the experts, then the universe is a 3-dimensional manifold which changes its shape and size with time due to the existence of matter, and it is a challenge to understand this evolution, too. Perhaps this evolution is driven by an energy which is, from the mathematical point of view, very similar to the elastic energy inside a surface, for example, a football. Perhaps this elastic energy is exactly the energy which is nowadays also used in image processing where moving surfaces also appear.

For esthetical reasons we would very much like to discuss the evolution of soap bubbles, believing that this model is relatively accessible. However, in order to simplify the presentation, we considers the evolution of curves instead of surfaces.

## Curves and curvature vector

A smooth curve is a subset $\Gamma \subseteq \mathbb{R}^{d}$ which is the image of a continuously differentiable function $u: I \rightarrow \mathbb{R}^{d}$ ( $I \subseteq \mathbb{R}$ an interval) with never vanishing derivative. The function $u$ is called parametrization of the curve $\Gamma$. One curve admits infinitely many parametrizations, since for every $C^{1}$ diffeomorphism $\theta: J \rightarrow I$ between two intervals $J, I \subseteq \mathbb{R}$ the composite function $u \circ \theta$ is also a parametrization.

Let $\Gamma$ be a curve with a $C^{2}$ parametrization $u: I \rightarrow \mathbb{R}^{d}$. Then, for every $x \in I$ the vector $u^{\prime}(x)$ is the tangent to $\Gamma$ at $u(x)$. The curvature vector at $u(x)$ is defined by

$$
\kappa(u(x)):=\frac{1}{\left|u^{\prime}(x)\right|}\left(\frac{u^{\prime}(x)}{\left|u^{\prime}(x)\right|}\right)^{\prime}
$$

where $|\cdot|$ denotes the euclidean norm in $\mathbb{R}^{d}$. The curvature vector $\kappa$ measures, by its length, how much the curve $\Gamma$ is curved (see Lemma 13.1 below). Moreover, for planar curves it points into the direction where the curve is convex. The curvature vector depends only on the point $u(x) \in \Gamma$ and shape of the curve $\Gamma$, but it does not depend on the particular parametrization of $\Gamma$. In fact, if $v: J \rightarrow \mathbb{R}^{d}$ is another $C^{2}$ parametrization of $\Gamma, v=u \circ \theta$ for some $C^{2}$ diffeomorphism $\theta: J \rightarrow I$, then, at the point $v(x)=u(\theta(x)) \in \Gamma$,

$$
\begin{aligned}
\left(\frac{1}{\left|v^{\prime}\right|}\left(\frac{v^{\prime}}{\left|v^{\prime}\right|}\right)^{\prime}\right)(x) & =\frac{1}{\left|(u \circ \theta)^{\prime}(x)\right|}\left(\frac{(u \circ \theta)^{\prime}(x)}{\left|(u \circ \theta)^{\prime}(x)\right|}\right)^{\prime} \\
& =\frac{1}{\left|u^{\prime}(\theta(x))\right|\left|\theta^{\prime}(x)\right|}\left(\frac{u^{\prime}(\theta(x))}{\left|u^{\prime}(\theta(x))\right|} \operatorname{sgn} \theta^{\prime}(x)\right)^{\prime} \\
& =\frac{1}{\left|u^{\prime}(\theta(x))\right|\left|\theta^{\prime}(x)\right|}\left(\frac{u^{\prime}}{\left|u^{\prime}\right|}\right)^{\prime}(\theta(x)) \theta^{\prime}(x) \operatorname{sgn} \theta^{\prime}(x) \\
& =\left(\frac{1}{\left|u^{\prime}\right|}\left(\frac{u^{\prime}}{\left|u^{\prime}\right|}\right)^{\prime}\right)(\theta(x)) .
\end{aligned}
$$

Given a $C^{2}$ parametrization $u: I \rightarrow \mathbb{R}^{d}$ of a curve $\Gamma, d s=\left|u^{\prime}(x)\right| d x$ is the infinitesimal arclength of the curve at $u(x), \int_{x}^{x+h}\left|u^{\prime}(\xi)\right| d \xi$ is the length of the arc segment running from $u(x)$ to $u(x+h)$ (respecting the orientation of $u$ ), and $\mathscr{E}(u)=\int_{I}\left|u^{\prime}(x)\right| d x$ is the total length of the curve $\Gamma$.

Lemma 13.1. Fix $x \in I$. For $h \in \mathbb{R}$ small, let $\theta(h)$ be the angle between $u^{\prime}(x+h)$ and $u^{\prime}(x)$, let $s(h)$ be the length of the arc segment running from $u(x)$ to $u(x+h)$. Then

$$
|\kappa(u(x))|=\lim _{h \rightarrow 0}\left|\frac{\theta(h)}{s(h)}\right| .
$$

Proof. Consider the triangle with vortices $0, \frac{u^{\prime}(x)}{\left|u^{\prime}(x)\right|}$ and $\frac{u^{\prime}(x+h)}{\left|u^{\prime}(x+h)\right|}$. By definition, $\theta(h)$ is the angle at the vortex 0 , and the two sides forming this angle have length equal to 1 . By bisecting the angle $\theta(h)$, one easily sees that

$$
\sin \frac{\theta(h)}{2}=\frac{1}{2}\left|\frac{u^{\prime}(x+h)}{\left|u^{\prime}(x+h)\right|}-\frac{u^{\prime}(x)}{\left|u^{\prime}(x)\right|}\right| .
$$

Then

$$
\begin{aligned}
\lim _{h \rightarrow 0}\left|\frac{\theta(h)}{s(h)}\right| & =\lim _{h \rightarrow 0}\left|\frac{\theta(h) / 2}{\sin (\theta(h) / 2)} \frac{2 \sin (\theta(h) / 2)}{s(h)}\right| \\
& =\lim _{h \rightarrow 0}\left|\left(\frac{u^{\prime}(x+h)}{\left|u^{\prime}(x+h)\right|}-\frac{u^{\prime}(x)}{\left|u^{\prime}(x)\right|}\right)\left(h \frac{1}{h} \int_{x}^{x+h}\left|u^{\prime}(\xi)\right| d \xi\right)^{-1}\right| \\
& =\left|\left(\frac{u^{\prime}(x)}{\left|u^{\prime}(x)\right|}\right)^{\prime} \frac{1}{\left|u^{\prime}(x)\right|}\right| \\
& =|\kappa(u(x))|
\end{aligned}
$$

Given a $C^{2}$ parametrization $u:[a, b] \rightarrow \mathbb{R}^{d}$ of a curve $\Gamma$, we define $s(x):=$ $\int_{a}^{x}\left|u^{\prime}(\xi)\right| d \xi$. Then $s:[a, b] \rightarrow[0, L]$ ( $L$ the length of the curve $\Gamma$ ) is strictly increasing, surjective and twice continuously differentiable. Moreover, $\gamma(x):=u\left(s^{-1}(x)\right)$ defines another $C^{2}$ parametrization of $\Gamma$, called the arclength parametrization. The arclength parametrization has unit speed, that is, the first derivative has unit length, and the second derivative coincides with the curvature vector. Moreover, with the help of the arclength parametrization one easily sees that the curvature vec-
tor is always orthogonal to the tangent vector. These properties are summarized in the following lemma.

Lemma 13.2. For every $x \in[0, L]$, the arclength parametrization $\gamma$ satisfies
a) $\left|\gamma^{\prime}(x)\right|=1$,
b) $\kappa(\gamma(x))=\gamma^{\prime \prime}(x)$, and
c) $\gamma^{\prime}(x) \gamma^{\prime \prime}(x)=0$.

Proof. One has,
(a) $\quad\left|\gamma^{\prime}(x)\right|=\left|u^{\prime}\left(s^{-1}(x)\right)\left(s^{-1}\right)^{\prime}(x)\right|=\frac{\left|u^{\prime}\left(s^{-1}(x)\right)\right|}{\left|s^{\prime}\left(s^{-1}(x)\right)\right|}=1$,
(b) $\quad \kappa(\gamma(x))=\frac{1}{\left|\gamma^{\prime}(x)\right|}\left(\frac{\gamma^{\prime}(x)}{\left|\gamma^{\prime}(x)\right|}\right)^{\prime}=\gamma^{\prime \prime}(x) \quad$ (by part (a)), and
(c) $\quad \gamma^{\prime}(x) \gamma^{\prime \prime}(x)=\frac{d}{d x} \frac{1}{2}\left|\gamma^{\prime}(x)\right|^{2}=0 \quad$ (by part (a)).

Example 13.3. Let $\Gamma$ be the circle with radius $r>0$ and center 0 . The function $u: \mathbb{R} \rightarrow \mathbb{R}^{2}, u(x)=(r \cos x, r \sin x)$ is a parametrization. For every $x \in \mathbb{R}$,

$$
\kappa(u(x))=-\frac{1}{r}(\cos x, \sin x)
$$

is the curvature vector at $u(x) \in \Gamma$. The curvature vector points into the disk enclosed by $\Gamma$ and has length $=1 / r$.

## The curve shortening flow

The curve shortening problem arises in several contexts and applications. Consider first the following problem: given two points in $\mathbb{R}^{d}$, find the shortest curve connecting the two points. If one gently ignores the well-known solution, that is, the straight line between the two points, then this problem - as it is posed - is exactly the problem of finding a minimizer of a real-valued function defined on a certain set of curves. This function assigns to each curve its length and is called the length functional in the following.

Similarly, one may consider the analogous problem of finding curves with minimal length for curves in Riemannian manifolds. A solution is then not so easy to imagine. By finding curves with minimal length, G. D. Birkhoff solved in 1917 a problem about the existence of periodic solutions of the Lagrange equations from classical mechanics; for this he was awarded the Bôcher Memorial Prize by the American Mathematical Society in 1923. Periodic solutions of the Lagrange equations correspond to closed curves of minimal length, that is, closed geodesics,
in certain Riemannian surfaces determined by the Lagrangian. Starting from appropriate closed curves, Birkhoff proposed a discrete curve shortening algorithm which, step by step, constructed closed curves of shorter and shorter length and he showed that the algorithm converges to a desired solution.


Fig. 13.1 Curve shortening by moving $B$ into the direction of the curvature vector

In $\mathbb{R}^{2}$, Birkhoff's idea of curve shortening can be described as follows. Let $\Gamma_{0}$ be an initial curve, and take three points $A, B$, $C \in \Gamma_{0}$ sufficiently close together and such that $B$ lies inside the arc segment AC. Then we may construct a shorter curve $\Gamma_{1}$ by simply taking the curve $\Gamma_{0}$ outside the arc segment AC, and by replacing the arc segment AC by the straight line $\overline{A C}$. This shortens the curve $\Gamma_{0}$ but the new curve $\Gamma_{1}$ is in general no longer parametrized by a $C^{1}$ function; $\Gamma_{1}$ is actually not really a curve in the sense of our definition. Intuitively, instead of replacing the arc segment AC by the line segment $\overline{A C}$, we may say that the point $B$ on the arc segment AC should be slightly moved into the direction of the curvature vector $\kappa(B)$. This should also shorten the curve since the curvature vector points into the direction where the straight line $\overline{A C}$ lies. Similarly, if $A_{0}, A_{1}, \ldots, A_{n}$ is a finite family of points on $\Gamma_{0}$ such that the open arc segments $A_{j} A_{j+1}$ are mutually disjoint, we may shorten the curve $\Gamma_{0}$, by taking for every $j=0, \ldots, n-1$ some point $B_{j} \in A_{j} A_{j+1}$ and by moving the curve slightly into the direction of the curvature vector $\kappa\left(B_{j}\right)$. By taking finer and finer partitions of the curve $\Gamma_{0}$, and by making shorter and shorter steps into the direction of the curvature vector, one may arrive at a continuous version of curve shortening. Without making a formal limiting process, it is conceivable that we are looking for a family of curves $\left(\Gamma_{t}\right)_{t \in[0, T]}$ parametrized by a function $u=u(t, x)$ (that is, $u(t, \cdot)$ is a parametrization of $\Gamma_{t}$ ) such that

$$
\begin{equation*}
u_{t}(t, x)=\kappa(u(t, x)) . \tag{13.1}
\end{equation*}
$$

Indeed, this equation expresses the idea that, at every time $t$, points of the curve $\Gamma_{t}$ moves into the direction of the curvature vector. In (13.1), the velocity $u_{t}$ with which the points move is equal to the curvature vector. Equation (13.1) is called the curve shortening flow equation. Depending on the particular context, it has to be equipped with boundary conditions and an initial condition.

Let us approach the curve shortening problem in a different way. Although the objects of interest are curves and although the length functional is defined on a set of curves, we think of it being defined on the corresponding set of all parametrizations. In this way, the length functional is defined on a subset of a linear space, and it is possible to study its continuity or differentiability, to consider gradients of the
length functional and to search for a minimizer by a steepest descent method, that is, by a gradient system. If $u$ is a solution of a gradient system associated with the length functional, then, for each "time" $t, u(t)$ is a parametrization of a curve $\Gamma_{t}$ and we expect that the length of $\Gamma_{t}$ is decreasing with time.

Instead of considering curves connecting two given points, let us consider in the following only closed curves, that is, by definition, curves which admit a periodic parametrization defined on $\mathbb{R}$. Without loss of generality, it suffices to consider $2 \pi$ periodic parametrizations, and we define accordingly

$$
\mathscr{P}_{2 \pi}:=\left\{u \in C^{1}\left(\mathbb{R} ; \mathbb{R}^{d}\right): u(x)=u(x+2 \pi) \text { and } u^{\prime}(x) \neq 0 \text { for every } x \in \mathbb{R}\right\} .
$$

This set is an open subset of the Banach space

$$
C_{2 \pi}^{1}:=\left\{u \in C^{1}\left(\mathbb{R} ; \mathbb{R}^{d}\right): u(x)=u(x+2 \pi) \text { for every } x \in \mathbb{R}\right\}
$$

Every element $u \in \mathscr{P}_{2 \pi}$ is a parametrization of some closed curve $\Gamma$ with finite length. In fact, $d s=\left|u^{\prime}(x)\right| d x$ is the infinitesimal arclength of the curve at $u(x)$, $\int_{a}^{b}\left|u^{\prime}(x)\right| d x$ is the length of the arc segment running from $u(a)$ to $u(b)$ (respecting the orientation of $u$ ), and

$$
\mathscr{E}(u)=\int_{0}^{2 \pi}\left|u^{\prime}(x)\right| d x
$$

is the total length of $\Gamma$. The function $\mathscr{E}: \mathscr{P}_{2 \pi} \rightarrow \mathbb{R}$ thus defined is the length functional on the set of closed curves. If $u, v \in \mathscr{P}_{2 \pi}$ are two parametrizations of the same curve $\Gamma, v=u \circ \theta$ for some diffeomorphism $\theta:[0,2 \pi] \rightarrow[0,2 \pi]$, then a change of variables implies

$$
\begin{aligned}
\mathscr{E}(v) & =\int_{0}^{2 \pi}\left|(u \circ \theta)^{\prime}(x)\right| d x \\
& =\int_{0}^{2 \pi}\left|u^{\prime}(\theta(x))\right|\left|\theta^{\prime}(x)\right| d x \\
& =\int_{0}^{2 \pi}\left|u^{\prime}(y)\right| d y \\
& =\mathscr{E}(u)
\end{aligned}
$$

Hence, the length functional $\mathscr{E}$ depends only on the curve $\Gamma$ and not on the particular parametrization $u \in \mathscr{P}_{2 \pi}$. Since the euclidean norm $|\cdot|$ is continuously differentiable in $\mathbb{R}^{d} \backslash\{0\}$, and since the derivative of elements in $\mathscr{P}_{2 \pi}$ is bounded away from 0 , it is straightforward to show that $\mathscr{E}$ is continuously differentiable on the set $\mathscr{P}_{2 \pi}$ (the $C^{1}$ topology is also crucial here), and that

$$
\mathscr{E}^{\prime}(u) \varphi=\int_{0}^{2 \pi} \frac{u^{\prime}(x) \varphi^{\prime}(x)}{\left|u^{\prime}(x)\right|} d x
$$

compare with Exercise 7.2. Given a parametrization $u \in \mathscr{P}_{2 \pi}$ of a curve $\Gamma$, it seems natural to consider the following inner product on

$$
L_{2 \pi}^{2}=\left\{u \in L_{l o c}^{2}\left(\mathbb{R} ; \mathbb{R}^{d}\right): u(x)=u(x+2 \pi) \text { for almost every } x \in \mathbb{R}\right\}
$$

namely the $L^{2}$ inner product with respect to arclength:

$$
\langle v, w\rangle_{g(u)}:=\int_{0}^{2 \pi} v(x) w(x)\left|u^{\prime}(x)\right| d x \quad\left(v, w \in L_{2 \pi}^{2}\right)
$$

In this way, we obtain a metric $g: \mathscr{P}_{2 \pi} \rightarrow \operatorname{Inner}\left(L_{2 \pi}^{2}\right)$. We calculate the gradient of $\mathscr{E}$ with respect to this metric $g$. First, let $u \in D\left(\nabla_{g} \mathscr{E}\right)$. By definition, there exists $v \in L_{2 \pi}^{2}, v=\nabla_{g} \mathscr{E}(u)$, such that, for every $\varphi \in C_{2 \pi}^{1}$,

$$
\begin{equation*}
\int_{0}^{2 \pi} \frac{u^{\prime}(x) \varphi^{\prime}(x)}{\left|u^{\prime}(x)\right|} d x=\mathscr{E}^{\prime}(u) \varphi=\langle v, \varphi\rangle_{g(u)}=\int_{0}^{2 \pi} v(x) \varphi(x)\left|u^{\prime}(x)\right| d x \tag{13.2}
\end{equation*}
$$

By definition of the Sobolev space $H_{2 \pi}^{1}$, this implies

$$
\frac{u^{\prime}}{\left|u^{\prime}\right|} \in H_{2 \pi}^{1} \text { and }\left(\frac{u^{\prime}}{\left|u^{\prime}\right|}\right)^{\prime}=-v\left|u^{\prime}\right|
$$

Conversely, let $u \in \mathscr{P}_{2 \pi}$ be such that $\frac{u^{\prime}}{\left|u^{\prime}\right|} \in H_{2 \pi}^{1}$. Then $\frac{1}{\left|u^{\prime}\right|}\left(\frac{u^{\prime}}{\left|u^{\prime}\right|}\right)^{\prime} \in L_{2 \pi}^{2}$, and the equation (13.2) holds with $v=\frac{1}{\left|u^{\prime}\right|}\left(\frac{u^{\prime}}{\left|u^{\prime}\right|}\right)^{\prime}$. Hence, $u \in D\left(\nabla_{g} \mathscr{E}\right)$, by the definition of the domain of the gradient. We have thus proved that

$$
\begin{aligned}
D\left(\nabla_{g} \mathscr{E}\right) & =\left\{u \in \mathscr{P}_{2 \pi}: \frac{u^{\prime}}{\left|u^{\prime}\right|} \in H_{2 \pi}^{1}\right\} \\
\nabla_{g} \mathscr{E}(u) & =-\frac{1}{\left|u^{\prime}\right|}\left(\frac{u^{\prime}}{\left|u^{\prime}\right|}\right)^{\prime}=-\kappa(u)
\end{aligned}
$$

where $\kappa(u)$ is the curvature vector defined in the previous section.
The gradient system associated with the length functional and the metric $g$ is thus exactly the curve shortening flow equation for closed curves:

$$
\begin{cases}u_{t}-\frac{1}{\left|u_{x}\right|}\left(\frac{u_{x}}{\left|u_{x}\right|}\right)_{x}=0 & \text { in }[0, T] \times \mathbb{R}  \tag{13.3}\\ u(t, x)=u(t, x+2 \pi) & \text { for }(t, x) \in[0, T] \times \mathbb{R} \\ u(0, x)=u_{0}(x) & \text { for } x \in \mathbb{R}\end{cases}
$$

Let us repeat that $u(t, \cdot) \in \mathscr{P}_{2 \pi}$ is a parametrization of a curve $\Gamma_{t}$, and a solution of this geometric evolution equation corresponds to a family $\left(\Gamma_{t}\right)_{t \in[0, T]}$ of curves. In particular, if $u$ and $v$ are solutions of the partial differential equation (13.3) such that, for every time $t, u(t, \cdot)$ and $v(t, \cdot)$ are parametrizations of the same curve $\Gamma_{t}$, then the two solutions are identified. We call a family $\left(\Gamma_{t}\right)_{t \in[0, T)}$ of curves a curve
shortening flow if it admits a parametrization $u=u(t, x)$ which is a maximal solution to the curve shortening flow equation (we skip the precise definition of solution, so one may think of classical solution).

Several questions arise in the context of problem (13.3): does one have existence and uniqueness of local or maximal solutions? For which initial curves? May we expect smooth, that is, $C^{\infty}$ soluions? What is the maximal existence time of a curve shortening flow? What is the behaviour near the maximal existence time? Do singularities develop, and if yes, which kind of singularities?

In order to get a first idea how answers may look like, the following example is helpful.

Example 13.4 (of a solution of the curve shortening flow equation). Let $\Gamma_{0}$ be the circle of radius $r_{0}>0$ and center 0 . We are looking for a solution family $\left(\Gamma_{t}\right)$ of the curve shortening flow equation. By reasons of symmetry, we make the ansatz that $\Gamma_{t}$ is a circle of radius $r(t)$ and center 0 . We look for a solution of the form

$$
u(t, x):=(r(t) \cos x, r(t) \sin x)
$$

Inserting this function into (13.3) and recalling Example 13.3, we see that $u$ is a solution of (13.3) if and only if the function $r$ is a solution of the ordinary differential equation

$$
\dot{r}+\frac{1}{r}=0, \quad r(0)=r_{0}
$$

Hence,

$$
u(t, x):=\sqrt{r_{0}^{2}-2 t}(\cos x, \sin x) \quad \text { for }(t, x) \in\left[0, \frac{1}{2} r_{0}^{2}\right) \times \mathbb{R}
$$

is indeed a solution of the curve shortening flow equation. This solution is maximal but not global. The maximal existence time $T_{\max }=\frac{1}{2} r_{0}^{2}$ is finite, and the solution shrinks to a point as time approaches the maximal existence time.

By using maximal regularity results in the Sobolev space $H_{2 \pi}^{2}$, by using fixed point theorems or the local inverse theorem, and by using regularity results for gradient systems, it is possible to prove the following local / maximal existence and uniqueness result. It is, however, not a straightforward consequence of Theorems 6.1 or 8.1 for abstract gradient systems.

Theorem 13.5. For every initial curve in $H_{2 \pi}^{2}$, the problem (13.3) admits a unique maximal solution $u \in C\left(\left[0, T_{\max }\right) ; H_{2 \pi}^{2}\right) \cap C^{\infty}\left(\left(0, T_{\max }\right) ; C_{2 \pi}^{\infty}\right)$.

More difficult is the question about the maximal existence time and the behaviour of the curve shortening flow near the maximal existence time. In order to give some short outlook, let us cite a few results in this context.

A closed, planar curve $\Gamma \subseteq \mathbb{R}^{2}$ is called a Jordan curve if it has no selfintersections, that is, if every parametrization $u \in \mathscr{P}_{2 \pi}$ of the curve is injective on $[0,2 \pi)$. By the Jordan curve theorem, every Jordan curve divides the plane into two connected components having the curve as boundary. One component is bounded, the other one is unbounded. We say that a Jordan curve is convex if the enclosed bounded component is convex. For convex initial curves and the curve shortening flow, the following can be said [Gage and Hamilton (1986)].

Theorem 13.6 (Gage \& Hamilton). Let $\left(\Gamma_{t}\right)$ be a curve shortening flow. If $\Gamma_{0}$ is a convex Jordan curve, then the $\Gamma_{t}$ are convex Jordan curves for every $t$ and they shrink to a point as time approaches the maximal existence time.

This result has been strengthened to general Jordan curves [Grayson (1987)].
Theorem 13.7 (Grayson). Let $\left(\Gamma_{t}\right)$ be a curve shortening flow. If $\Gamma_{0}$ is a Jordan curve, then the $\Gamma_{t}$ are Jordan curves for every $t$. Moreover, $\Gamma_{t_{0}}$ is convex for some $t_{0}$ and $\Gamma_{t}$ shrinks to a point as time approaches the maximal existence time.

Variants of the length functional may be considered on Riemannian manifolds, for example on the space $\mathbb{R}^{d}$ equipped with a metric $g$. The length of a closed curve is in this case

$$
\mathscr{E}(u)=\int_{0}^{2 \pi}\left|u^{\prime}(x)\right|_{g(u(x))} d x
$$

Clearly, minimizers of this functional, defined on the set of all curves connecting two given points, need no longer be straight lines. For closed curves, solutions of the associated curve shortening equation (the associated $L_{2 \pi}^{2}$ gradient system) need no longer shrink to points. In the Lagrange equations from classical mechanics, the above more general situation appears very naturally, and we refer back to [Birkhoff (1917)] or [Cartan (1967), pp. 287-].

## References

[Absil et al. (2005)] Absil, P.-A., Mahony, R., Andrews, B. : Convergence of the iterates of descent methods for analytic cost functions. SIAM J. Optim. 16, 2005, 531-547 (electronic).
[Adams (1975)] Adams, R. A. : Sobolev Spaces. Academic Press, New York, 1975.
[Adams and Fournier (2003)] Adams, R. A., Fournier, J. J. F. : Sobolev Spaces. Vol. 140 of Pure and applied mathematics. Academic Press, Amsterdam, 2003.
[Amann (1995)] Amann, H. : Linear and Quasilinear Parabolic Problems: Abstract Linear Theory. Vol. 89 of Monographs in Mathematics. Birkhäuser Verlag, Basel, 1995.
[Ambrosio (2008)] Ambrosio, L. : Gradient flows in metric spaces and in the space of probability measures, and applications to Fokker-Planck equations with respect to log-concave measures. Bolletino della Unione Matematica Italiana IX 1, 2008, 223-240.
[Ambrosio et al. (2005)] Ambrosio, L., Gigli, N., Savaré, G. : Gradient Flows. Lectures in Mathematics ETH Zürich. Birkhäuser, Basel, 2005.
[Angenent (1990a)] Angenent, S. : Parabolic equations for curves on surfaces. I. Curves with pintegrable curvature. Ann. of Math. (2) 132 (3), 1990a, 451-483. URL http://dx.doi.org/10.2307/1971426
[Angenent (1990b)] Angenent, S. B. : Nonlinear analytic semiflows. Proc. Roy. Soc. Edinburgh 115A, 1990b, 91-107.
[Arendt (2004)] Arendt, W., 2004. : Semigroups and evolution equations: functional calculus, regularity and kernel estimates. In: Handbook of Differential Equations (C. M. Dafermos, E. Feireisl eds.). Elsevier/North Holland, pp. 1-85.
[Aubin (1979)] Aubin, J. : Applied Functional Analysis. Pure Applied Mathematics. John Wiley \& Sons, New York, Chichester, Brisbane, Toronto, 1979.
[Birkhoff (1917)] Birkhoff, G. D. : Dynamical systems with two degrees of freedom. Trans. Amer. Math. Soc. 18, 1917, 199-300.
[Boys (1896)] Boys, C. V. : Soap Bubbles and the Forces which Mould Them. E. \& J. B. Young \& Co., New York, 1896.
[Brézis (1973)] Brézis, H. : Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert. Vol. 5 of North Holland Mathematics Studies. North-Holland, Amsterdam, London, 1973.
[Brézis (1992)] Brézis, H. : Analyse fonctionnelle. Masson, Paris, 1992.
[Cartan (1967)] Cartan, H. : Calcul différentiel. Hermann, Paris, 1967.
[Cipriani and Grillo (2003)] Cipriani, F., Grillo, G. : Nonlinear Markov semigroups, nonlinear Dirichlet forms and application to minimal surfaces. J. reine angew. Math. 562, 2003, 201235.
[Clément and Li (1994)] Clément, P., Li, S. : Abstract parabolic quasilinear problems and application to a groundwater flow problem. Adv. Math. Sci. Appl. 3, 1994, 17-32.
[Curry (1944)] Curry, H. B. : The method of steepest descent for non-linear minimization problems. Quart. Appl. Math. 2, 1944.
[Dautray and Lions (1985)] Dautray, R., Lions, J. : Analyse mathématique et calcul numérique pour les sciences et les techniques. Vol. I. INSTN: Collection Enseignement. Masson, Paris, 1985.
[Dautray and Lions (1987)] Dautray, R., Lions, J. : Analyse mathématique et calcul numérique pour les sciences et les techniques. Vol. VIII. INSTN: Collection Enseignement. Masson, Paris, 1987.
[Dautray and Lions (1988)] Dautray, R., Lions, J. : Analyse mathématique et calcul numérique pour les sciences et les techniques. Vol. VI. INSTN: Collection Enseignement. Masson, Paris, 1988.
[De Giorgi et al. (1980)] De Giorgi, E., Marino, A., Tosques, M. : Problems of evolution in metric spaces and maximal decreasing curve. Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8) 68 (3), 1980, 180-187.
[Denk et al. (2003)] Denk, R., Hieber, M., Prüss, J. : $\mathscr{R}$-Boundedness, Fourier Multipliers and Problems of Elliptic and Parabolic Type. Vol. 166 of Memoirs Amer. Math. Soc. Amer. Math. Soc., Providence, R.I., 2003.
[Diestel and Uhl (1977)] Diestel, J., Uhl, J. J. : Vector Measures. Vol. 15 of Mathematical Surveys. Amer. Math. Soc., Providence, R.I., 1977.
[Dore (1993)] Dore, G., 1993. : L ${ }^{p}$ regularity for abstract differential equations. In: Functional analysis and related topics, 1991 (Kyoto). Vol. 1540 of Lecture Notes in Math. Springer, Berlin, pp. 25-38.
[Drábek and Milota (2007)] Drábek, P., Milota, J. : Methods of nonlinear analysis. Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks]. Birkhäuser Verlag, Basel, 2007.
[Escher et al. (2003)] Escher, J., Prüss, J., Simonett, G., 2003. : A new approach to the regularity of solutions of parabolic equations. In: Evolution Equations: Proceedings in Honor of J. A. Goldstein's 60th Birthday. Marcel Dekker, New York, pp. 167-190.
[Evans (1998)] Evans, L. C. : Partial Differential Equations. Vol. 19 of Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 1998.
[Feireisl and Simondon (2000)] Feireisl, E., Simondon, F. : Convergence for semilinear degenerate parabolic equations in several space dimensions. J. Dynam. Differential Equations 12, 2000, 647-673.
[Gage and Hamilton (1986)] Gage, M., Hamilton, R. S. : The heat equation shrinking convex plane curves. J. Differential Geom. 23 (1), 1986, 69-96.
URL http://projecteuclid.org/getRecord?id=euclid.jdg/ 1214439902
[Gilbarg and Trudinger (2001)] Gilbarg, D., Trudinger, N. S. : Elliptic Partial Differential Equations of Second Order. Springer Verlag, Berlin, Heidelberg, New York, 2001.
[Goldstein (1962)] Goldstein, A. A. : Cauchy's method of minimization. Numer. Math. 4, 1962, 146-150.
[Grayson (1987)] Grayson, M. A. : The heat equation shrinks embedded plane curves to round points. J. Differential Geom. 26 (2), 1987, 285-314.
URL http://projecteuclid.org/getRecord?id=euclid.jdg/ 1214441371
[Haraux (1990)] Haraux, A. : Systèmes dynamiques dissipatifs et applications. Masson, Paris, 1990.
[Hille and Phillips (1957)] Hille, E., Phillips, R. S. : Functional Analysis and Semi-Groups. Amer. Math. Soc., Providence, R.I., 1957.
[Huisken and Polden (1999)] Huisken, G., Polden, A., 1999. : Geometric evolution equations for hypersurfaces. In: Calculus of variations and geometric evolution problems (Cetraro, 1996). Vol. 1713 of Lecture Notes in Math. Springer, Berlin, pp. 45-84.
[Jahnke (2003)] Jahnke, H. N. (Ed.). : A history of analysis. Vol. 24 of History of Mathematics. American Mathematical Society, Providence, RI, 2003.
[Kantorovič (1947)] Kantorovič, L. V. : On the method of steepest descent. Doklady Akad. Nauk SSSR (N. S.) 56, 1947, 233-236.
[Kunstmann and Weis (2004)] Kunstmann, P. C., Weis, L., 2004. : Maximal $L^{p}$ regularity for parabolic equations, Fourier multiplier theorems and $H^{\infty}$ functional calculus. In: Levico Lectures, Proceedings of the Autumn School on Evolution Equations and Semigroups (M. Iannelli, R. Nagel, S. Piazzera eds.). Vol. 69. Springer Verlag, Heidelberg, Berlin, pp. 65-320.
[Ladyženskaja et al. (1967)] Ladyženskaja, O. A., Solonnikov, V. A., Ural'ceva, N. N. : Linear and quasilinear equations of parabolic type. Translated from the Russian by S. Smith. Translations of Mathematical Monographs, Vol. 23. American Mathematical Society, Providence, R.I., 1967.
[Lions (1969)] Lions, J. : Quelques méthodes de résolution des problèmes aux limites non linéaires. Dunod, Gauthier-Villars, Paris, 1969.
[Łojasiewicz (1963)] Łojasiewicz, S., 1963. : Une propriété topologique des sous-ensembles analytiques réels. In: Colloques internationaux du C.N.R.S.: Les équations aux dérivées partielles, Paris (1962). Editions du C.N.R.S., Paris, pp. 87-89.
[Łojasiewicz (1965)] Łojasiewicz, S. : Ensembles semi-analytiques, 1965, preprint, I.H.E.S. Bures-sur-Yvette.
[Lunardi (1995)] Lunardi, A. : Analytic Semigroups and Optimal Regularity in Parabolic Problems. Vol. 16 of Progress in Nonlinear Differential Equations and Their Applications. Birkhäuser, Basel, 1995.
[Lunardi (2009)] Lunardi, A. : Interpolation theory, 2nd Edition. Appunti. Scuola Normale Superiore di Pisa (Nuova Serie). [Lecture Notes. Scuola Normale Superiore di Pisa (New Series)]. Edizioni della Normale, Pisa, 2009.
[Neuberger (1997)] Neuberger, J. W. : Sobolev Gradients and Differential Equations. Vol. 1670 of Lect. Notes Math. Springer Verlag, Berlin, Heidelberg, New York, 1997.
[Neuberger (2005)] Neuberger, J. W. : Prospects for a central theory of partial differential equations. Math. Intelligencer 27, 2005, 47-55.
[Ouhabaz (2004)] Ouhabaz, E. M. : Analysis of Heat Equations on Domains. Vol. 30 of London Mathematical Society Monographs. Princeton University Press, Princeton, 2004.
[Palis and de Melo (1982)] Palis, J., de Melo, W. : Geometric Theory of Dynamical Systems. An Introduction. Springer Verlag, New York, Heidelberg, Berlin, 1982.
[Perelman and Petrunin (1994)] Perelman, G., Petrunin, A. : Quasigeodesics and gradient curves in Alexandrov spaces, 1994, unpublished preprint.
[Perona and Malik (1990)] Perona, P., Malik, J. : Scale-Space and Edge Detection Using Anisotropic Diffusion. IEEE Transactions on Pattern Analysis and Machine Intelligence 12 (7), 1990, 629-639.
[Simon (1983)] Simon, L. : Asymptotics for a class of non-linear evolution equations, with applications to geometric problems. Ann. of Math. 118, 1983, 525-571.
[Simon (1996)] Simon, L. : Theorems on Regularity and Singularity of Energy Minimizing Maps. Lecture in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 1996.
[Temam (1984)] Temam, R. : Navier-Stokes Equations. Vol. 2 of Studies in Mathematics and its Applications. Elsevier Science Publishers, 1984.
[Triebel (1983)] Triebel, H. : Theory of Function Spaces. Birkhäuser, Basel, 1983.
[Whiteside (ed.) (1968)] Whiteside (ed.), D. T. : The Mathematical Papers of Isaac Newton, Volume II, 1667-1670. Cambridge University Press, Cambridge, 1968.
[Whiteside (ed.) (1969)] Whiteside (ed.), D. T. : The Mathematical Papers of Isaac Newton, Volume III, 1670-1673. Cambridge University Press, Cambridge, 1969.
[Wolfe (1969)] Wolfe, P. : Convergence conditions for ascent methods. SIAM Rev. 11, 1969, 226235.
[Zeidler (1990)] Zeidler, E. : Nonlinear Functional Analysis and Its Applications. I. Springer Verlag, New York, Berlin, Heidelberg, 1990.

## Appendix A

## Primer on topology

It is the purpose of this introductory chapter to recall some basic facts about metric spaces, sequences in metric spaces, compact metric spaces, and continuous functions between metric spaces. Most of the material should be known, and if it is not known in the context of metric spaces, it has certainly been introduced on $\mathbb{R}^{d}$. The generalization to metric spaces should be straightforward, but it is nevertheless worthwhile to spend some time on the examples.

We also introduce some further notions from topology which may be new; see for example the definitions of density or of completion of a metric space.

## A. 1 Metric spaces

Definition A.1. Let $M$ be a set. We call a function $d: M \times M \rightarrow \mathbb{R}_{+}$a metric or a distance on $M$ if for every $x, y, z \in M$
(i) $d(x, y)=0$ if and only if $x=y$,
(ii) $d(x, y)=d(y, x)$ (symmetry), and
(iii) $d(x, y) \leq d(x, z)+d(z, y)$ (triangle inequality).

A pair $(M, d)$ of a set $M$ and a metric $d$ on $M$ is called a metric space.
It will be convenient to write only $M$ instead of $(M, d)$ if the metric $d$ on $M$ is known from the context, and to speak of a metric space $M$.

Example A.2. 1. Let $M \subseteq \mathbb{R}^{d}$ and

$$
d(x, y):=\sum_{i=1}^{d}\left|x_{i}-y_{i}\right|
$$

or

$$
d(x, y):=\left(\sum_{i=1}^{d}\left|x_{i}-y_{i}\right|^{2}\right)^{\frac{1}{2}}
$$

Then $(M, d)$ is a metric space. The second metric is called the euclidean metric. Often, if the metric on $\mathbb{R}^{d}$ is not explicitly given, we mean the euclidean metric.
2. Let $M \subseteq C([0,1])$, the space of all continuous functions on the interval $[0,1]$, and

$$
d(f, g):=\sup _{x \in[0,1]}|f(x)-g(x)| .
$$

Then $(M, d)$ is a metric space.
3. Let $M$ be any set and

$$
d(x, y):= \begin{cases}0 & \text { if } x=y \\ 1 & \text { otherwise }\end{cases}
$$

Then $(M, d)$ is a metric space. The metric $d$ is called the discrete metric.
4. Let $(M, d)$ be a metric space. Then

$$
d_{1}(x, y):=\frac{d(x, y)}{1+d(x, y)}
$$

and

$$
d_{2}(x, y):=\min \{d(x, y), 1\}
$$

define also metrics on $M$.
5. Let $M=C(\mathbb{R})$, the space of all continuous functions on $\mathbb{R}$, and let

$$
d_{n}(f, g):=\sup _{x \in[-n, n]}|f(x)-g(x)| \quad(n \in \mathbb{N})
$$

and

$$
d(f, g):=\sum_{n \in \mathbb{N}} 2^{-n} \frac{d_{n}(f, g)}{1+d_{n}(f, g)}
$$

Then $(M, d)$ is a metric space. Note that the functions $d_{n}$ are not metrics for any $n \in \mathbb{N}$ !
6. Let $(M, d)$ be a metric space. Then any subset $\tilde{M} \subseteq M$ is a metric space for the induced metric

$$
\tilde{d}(x, y)=d(x, y), \quad x, y \in \tilde{M}
$$

We may sometimes say that $\tilde{M}$ is a subspace of $M$, that is, a subset and a metric space, but certainly this is not to be understood in the sense of linear subspaces of vector spaces ( $M$ need not be a vector space).
7. Let $\left(M_{n}, d_{n}\right)$ be metric spaces $(n \in \mathbb{N})$. Then the cartesian product $M:=\bigotimes_{n \in \mathbb{N}} M_{n}$ is a metric space for the metric

$$
d(x, y):=\sum_{n \in \mathbb{N}} 2^{-n} \min \left\{d_{n}\left(x_{n}, y_{n}\right), 1\right\} .
$$

Clearly, in a similar way, every finite cartesian product of metric spaces is a metric space.

Definition A.3. Let $(M, d)$ be a metric space.
a) For every $x \in M$ and every $r>0$ we define the open ball $B(x, r):=\{y \in M$ : $d(x, y)<r\}$ with center $x$ and radius $r$.
b) A set $O \subseteq M$ is called open if for every $x \in O$ there exists some $r>0$ such that $B(x, r) \subseteq O$.
c) A set $A \subseteq M$ is called closed if its complement $A^{c}=M \backslash A$ is open.
d) A set $U \subseteq M$ is called a neighbourhood of $x \in M$ if there exists $r>0$ such that $B(x, r) \subseteq U$.

Remark A.4. (a) The notions open, closed, neighbourhood depend on the set $M!!$ For example, $M$ is always closed and open in $M$. The set $\mathbb{Q}$ is not closed in $\mathbb{R}$ (for the euclidean metric), but it is closed in $\mathbb{Q}$ for the induced metric! Therefore, one should always say in which metric space some given set is open or closed.
(b) Clearly, a set $O \subseteq M$ is open (in $M$ ) if and only if it is a neighbourhood of every of its elements.

Lemma A.5. Let $(M, d)$ be a metric space. The following are true:
a) Arbitrary unions of open sets are open. That means: if $\left(O_{i}\right)_{i \in I}$ is an arbitrary family of open sets (no restrictions on the index set I), then $\bigcup_{i \in I} O_{i}$ is open.
b) Arbitrary intersections of closed sets are closed. That means: if $\left(A_{i}\right)_{i \in I}$ is an arbitrary family of closed sets, then $\bigcap_{i \in I} A_{i}$ is closed.
c) Finite intersections of open sets are open.
d) Finite unions of closed sets are closed.

Proof. (a) Let $\left(O_{i}\right)_{i \in I}$ be an arbitrary family of open sets and let $O:=\bigcup_{i \in I} O_{i}$. If $x \in O$, then $x \in O_{i}$ for some $i \in I$, and since $O_{i}$ is open, $B(x, r) \subseteq O_{i}$ for some $r>0$. This implies that $B(x, r) \subseteq O$, and therefore $O$ is open.
(c) Next let $\left(O_{i}\right)_{i \in I}$ be a finite family of open sets and let $O:=\bigcap_{i \in I} O_{i}$. If $x \in O$, then $x \in O_{i}$ for every $i \in I$. Since the $O_{i}$ are open, there exist $r_{i}$ such that $B\left(x, r_{i}\right) \subseteq O_{i}$. Let $r:=\min _{i \in I} r_{i}$ which is positive since $I$ is finite. By construction, $B(x, r) \subseteq O_{i}$ for every $i \in I$, and therefore $B(x, r) \subseteq O$, that is, $O$ is open.

The proofs for closed sets are similar or follow just from the definition of closed sets and the above two assertions.

Exercise A. 6 Determine all open sets (respectively, all closed sets) of a metric space $(M, d)$, where $d$ is the discrete metric.

Exercise A. 7 Show that a ball $B(x, r)$ in a metric space $M$ is always open. Show also that

$$
\bar{B}(x, r):=\{y \in M: d(x, y) \leq r\}
$$

is always closed.

Definition A. 8 (Closure, interior, boundary). Let $(M, d)$ be a metric space and let $S \subseteq M$ be a subset. Then the set $\bar{S}:=\bigcap\{A: A \subseteq M$ is closed and $S \subseteq A\}$ is called the closure of $S$. The set $S^{\circ}:=\bigcup\{O: O \subseteq M$ is open and $O \subseteq S\}$ is called the interior of $S$. Finally, we call $\partial S:=\left\{x \in M: \forall \varepsilon>0 B(x, \varepsilon) \cap S \neq \emptyset\right.$ and $\left.B(x, \varepsilon) \cap S^{c} \neq \emptyset\right\}$ the boundary of $S$.

By Lemma A.5, the closure of a set $S$ is always closed (arbitrary intersections of closed sets are closed). By definition, $\bar{S}$ is the smallest closed set which contains $S$. Similarly, the interior of a set $S$ is always open, and by definition it is the largest open set which is contained in $S$. Note that the interior might be empty.

Exercise A. 9 Give an example of a metric space $M$ and some $x \in M, r>0$, to show that $\bar{B}(x, r)$ need not coincide with the closure of $B(x, r)$.

Exercise A.10 Let $(M, d)$ be a metric space and consider the metrics $d_{1}$ and $d_{2}$ from Example A. 2 (4). Show that the set of all open subsets, closed subsets or neighbourhoods of $M$ is the same for the three given metrics.

The set of all open subsets is also called the topology of M. The three metrics $d$, $d_{1}$ and $d_{2}$ thus induce the same topology. Sometimes it is good to know that one can pass from a given metric $d$ to a finite metric ( $d_{1}$ and $d_{2}$ take only values between 0 and 1) without changing the topology.

## A. 2 Sequences, convergence

Throughout the following, sequences will be denoted by $\left(x_{n}\right)$. Only when it is necessary, we make precise the index $n$; usually, $n \geq 0$ or $n \geq 1$, but sometimes we will also consider finite sequences or sequences indexed by $\mathbb{Z}$.

Definition A.11. Let $(M, d)$ be a metric space.
a) We call a sequence $\left(x_{n}\right) \subseteq M$ a Cauchy sequence if for every $\varepsilon>0$ there exists $n_{0}$ such that for every $n, m \geq n_{0}$ one has $d\left(x_{n}, x_{m}\right)<\varepsilon$.
b) We say that a sequence $\left(x_{n}\right) \subseteq M$ converges to some element $x \in M$ if for every $\varepsilon>0$ there exists $n_{0}$ such that for every $n \geq n_{0}$ one has $d\left(x_{n}, x\right)<\varepsilon$. If $\left(x_{n}\right)$ converges to $x$, we also write $\lim _{n \rightarrow \infty} x_{n}=x$ or $x_{n} \rightarrow x$ as $n \rightarrow \infty$.

Exercise A. 12 Let $C([0,1])$ be the metric space from Example A. 2 (2). Show that a sequence $\left(f_{n}\right) \subseteq C([0,1])$ converges to some $f$ for the metric $d$ if and only if it converges uniformly. We say that the metric $d$ induces the topology of uniform convergence.

Show also that a sequence $\left(f_{n}\right) \subseteq C(\mathbb{R})$ (Example A. 2 (5)) converges to some $f$ for the metric $d$ if and only if it converges uniformly on compact subsets of $\mathbb{R}$. In this example, we say that the metric $d$ induces the topology of local uniform convergence.

Exercise A.13 Determine all Cauchy sequences and all convergent sequences in a discrete metric space.
Lemma A.14. Let $M$ be a metric space and $\left(x_{n}\right) \subseteq M$ be a sequence. Then:
a) $\lim _{n \rightarrow \infty} x_{n}=x$ for some element $x \in M$ if and only if for every neighbourhood $U$ of $x$ there exists $n_{0}$ such that for every $n \geq n_{0}$ one has $x_{n} \in U$.
b) (Uniqueness of the limit) If $\lim _{n \rightarrow \infty} x_{n}=x$ and $\lim _{n \rightarrow \infty} x_{n}=y$, then $x=y$.

Lemma A.15. A set $A \subseteq M$ is closed if and only iffor every sequence $\left(x_{n}\right) \subseteq A$ which converges to some $x \in M$ one has $x \in A$.

Proof. Assume first that $A$ is closed and let $\left(x_{n}\right) \subseteq A$ be convergent to $x \in M$. If $x$ does not belong to $A$, then it belongs to $A^{c}$ which is open. By definition, there exists $\varepsilon>0$ such that $B(x, \varepsilon) \subseteq A^{c}$. Given this $\varepsilon$, there exists $n_{0}$ such that $x_{n} \in B(x, \varepsilon)$ for every $n \geq n_{0}$, a contradiction to the assumption that $x_{n} \in A$. Hence, $x \in A$.

On the other hand, assume that $\lim _{n \rightarrow \infty} x_{n}=x \in A$ for every convergent $\left(x_{n}\right) \subseteq A$ and assume in addition that $A$ is not closed or, equivalently, that $A^{c}$ is not open. Then there exists $x \in A^{c}$ such that for every $n \in \mathbb{N}$ the set $B\left(x, \frac{1}{n}\right) \cap A$ is nonempty. From this one can construct a sequence $\left(x_{n}\right) \subseteq A$ which converges to $x$, which is a contradiction because $x \in A^{c}$.

Lemma A.16. Let $(M, d)$ be a metric space, and let $S \subseteq M$ be a subset. Then

$$
\begin{aligned}
\bar{S} & =\left\{x \in M: \exists\left(x_{n}\right) \subseteq S \text { s.t. } \lim _{n \rightarrow \infty} x_{n}=x\right\} \\
& =\left\{x \in M: d(x, S):=\inf _{y \in S} d(x, y)=0\right\} .
\end{aligned}
$$

Proof. Let

$$
A:=\left\{x \in M: \exists\left(x_{n}\right) \subseteq S \text { s.t. } \lim _{n \rightarrow \infty} x_{n}=x\right\}
$$

and

$$
B:=\left\{x \in M: d(x, S):=\inf _{y \in S} d(x, y)=0\right\} .
$$

These two sets are clearly equal by the definition of the inf and the definition of convergence. Moreover, the set $B$ is closed by the following argument. Assume that $\left(x_{n}\right) \subseteq B$ is convergent to $x \in M$. By definition of $B$, for every $n$ there exists $y \in S$ such that $d\left(x_{n}, y_{n}\right) \leq 1 / n$. Hence,

$$
\limsup _{n \rightarrow \infty} d\left(x, y_{n}\right) \leq \limsup _{n \rightarrow \infty} d\left(x, x_{n}\right)+\limsup _{n \rightarrow \infty} d\left(x_{n}, y_{n}\right)=0
$$

so that $x \in B$.
Clearly, $B$ contains $S$, and since $B$ is closed, $B$ contains $\bar{S}$. It remains to show that $B \subseteq \bar{S}$. If this is not true, then there exists $x \in B \backslash \bar{S}$. Since the complement of $\bar{S}$ is open in $M$, there exists $r>0$ such that $B(x, r) \cap \bar{S}=\emptyset$, a contradiction to the definition of $B$.

Definition A.17. A metric space $(M, d)$ is called complete if every Cauchy sequence converges.

Exercise A. 18 Show that the spaces $\mathbb{R}^{d}, C([0,1])$ and $C(\mathbb{R})$ are complete. Show also that any discrete metric space is complete.

Lemma A.19. A subspace $N \subseteq M$ of a complete metric space is complete if and only if it is closed in $M$.
Proof. Assume that $N \subseteq M$ is closed, and let $\left(x_{n}\right)$ be a Cauchy sequence in $N$. By the assumption that $M$ is complete, $\left(x_{n}\right)$ is convergent to some element $x \in M$. Since $N$ is closed, $x \in N$.

Assume on the other hand that $N$ is complete, and let $\left(x_{n}\right) \subseteq N$ be convergent to some element $x \in M$. Clearly, every convergent sequence is also a Cauchy sequence, and since $N$ is complete, $\left(x_{n}\right)$ converges to some element $y \in N$. By uniqueness of the limit, $x=y \in N$. Hence, $N$ is closed.

## A. 3 Compactness

Definition A.20. We say that a metric space $(M, d)$ is compact if for every open covering there exists a finite subcovering, that is, whenever $\left(O_{i}\right)_{i \in I}$ is a family of open sets (no restrictions on the index set $I$ ) such that $M=\bigcup_{i \in I} O_{i}$, then there exists a finite subset $I_{0} \subseteq I$ such that $M=\bigcup_{i \in I_{0}} O_{i}$.

Lemma A.21. A metric space $(M, d)$ is compact if and only if it is sequentially compact, that is, if and only if every sequence $\left(x_{n}\right) \subseteq M$ has a convergent subsequence.

Proof. Assume that $M$ is compact and let $\left(x_{n}\right) \subseteq M$. Assume that $\left(x_{n}\right)$ does not have a convergent subsequence. Then for every $x \in M$ there exists $\varepsilon_{x}>0$ such that $B\left(x, \varepsilon_{x}\right)$ contains only finitely many elements of $\left\{x_{n}\right\}$. Note that $\left(B\left(x, \varepsilon_{x}\right)\right)_{x \in M}$ is an open covering of $M$ so that by the compactness of $M$ there exists a finite subset $N \subseteq$ $M$ such that $M=\bigcup_{x \in N} B\left(x, \varepsilon_{x}\right)$. But this means that $\left(x_{n}\right)$ takes only finitely many values, and hence there exists even a constant subsequence which is in particular also convergent; a contradiction to the assumption on $\left(x_{n}\right)$.

On the other hand, assume that $M$ is sequentially compact and let $\left(O_{i}\right)_{i \in I}$ be an open covering of $M$. We first show that there exists $\varepsilon>0$ such that for every $x \in M$ there exists $i_{x} \in I$ with $B(x, \varepsilon) \subseteq O_{i_{x}}$. If this were not true, then for every $n \in \mathbb{N}$ there exists $x_{n}$ such that $B\left(x_{n}, \frac{1}{n}\right) \nsubseteq O_{i}$ for every $i \in I$. Passing to a subsequence, we may assume that $\left(x_{n}\right)$ is convergent to some $x \in M$. There exists some $i_{0} \in I$ such that $x \in O_{i_{0}}$, and since $O_{i_{0}}$ is open, we find some $\varepsilon>0$ such that $B(x, \varepsilon) \subseteq O_{i_{0}}$. Let $n_{0}$ be such that $\frac{1}{n_{0}}<\frac{\varepsilon}{2}$. By the triangle inequality, for every $n \geq n_{0}$ we have $B\left(x_{n}, \frac{1}{n}\right) \subseteq B(x, \varepsilon) \subseteq O_{i_{0}}$, a contradiction to the construction of the sequence $\left(x_{n}\right)$.

Next we show that $M=\bigcup_{j=1}^{n} B\left(x_{j}, \varepsilon\right)$ for a finite family of $x_{j} \in M$. Choose any $x_{1} \in M$. If $B\left(x_{1}, \varepsilon\right)=M$, then we are already done. Otherwise we find $x_{2} \in M \backslash$ $B\left(x_{1}, \varepsilon\right)$. If $B\left(x_{1}, \varepsilon\right) \cup B\left(x_{2}, \varepsilon\right) \neq M$, then we even find $x_{3} \in M$ which does not belong to $B\left(x_{1}, \varepsilon\right) \cup B\left(x_{2}, \varepsilon\right)$, and so on. If $\bigcup_{j=1}^{n} B\left(x_{j}, \varepsilon\right)$ is never all of $M$, then we find actually a sequence $\left(x_{j}\right)$ such that $d\left(x_{j}, x_{k}\right) \geq \varepsilon$ for all $j \neq k$. This sequence can not have a convergent subsequence, a contradiction to sequential compactness.

Since every of the $B\left(x_{j}, \varepsilon\right)$ is a subset of $O_{i_{x_{j}}}$ for some $i_{x_{j}} \in I$, we have proved that $M=\bigcup_{j=1}^{n} O_{i_{x_{j}}}$, i.e. the open covering $\left(O_{i}\right)$ admits a finite subcovering. The proof is complete.

Lemma A.22. Any compact metric space is complete.
Proof. Let $\left(x_{n}\right)$ be a Cauchy sequence in $M$. By the preceeding lemma, there exists a subsequence which converges to some $x \in M$. If a subsequence of a Cauchy sequence converges, then the sequence itself converges, too.

## A. 4 Continuity

Definition A.23. Let $\left(M_{1}, d_{1}\right),\left(M_{2}, d_{2}\right)$ be two metric spaces, and let $f: M_{1} \rightarrow M_{2}$ be a function.
a) We say that $f$ is continuous at some point $x \in M_{1}$ if

$$
\forall \varepsilon>0 \exists \delta>0 \forall y \in B(x, \delta): d_{2}(f(x), f(y))<\varepsilon .
$$

b) We say that $f$ is continuous if it is continuous at every point.
c) We say that $f$ is uniformly continuous if

$$
\forall \varepsilon>0 \exists \delta>0 \forall x, y \in M_{1}: d_{1}(x, y)<\delta \Rightarrow d_{2}(f(x), f(y))<\varepsilon .
$$

d) We say that $f$ is Lipschitz continuous if

$$
\exists L \geq 0 \forall x, y \in M: d_{2}(f(x), f(y)) \leq L d_{1}(x, y) .
$$

Lemma A.24. A function $f: M_{1} \rightarrow M_{2}$ between two metric spaces is continuous at some point $x \in M_{1}$ if and only if it is sequentially continuous at $x$, that is, if and only iffor every sequence $\left(x_{n}\right) \subseteq M_{1}$ which converges to $x$ one has $\lim _{n \rightarrow \infty} f\left(x_{n}\right)=f(x)$.
Proof. Assume that $f$ is continuous at $x \in M_{1}$ and let $\left(x_{n}\right)$ be convergent to $x$. Let $\varepsilon>0$. There exists $\delta>0$ such that for every $y \in B(x, \delta)$ one has $f(y) \in B(f(x), \varepsilon)$. By definition of convergence, there exists $n_{0}$ such that for every $n \geq n_{0}$ one has $x_{n} \in B(x, \delta)$. For this $n_{0}$ and every $n \geq n_{0}$ one has $f\left(x_{n}\right) \in B(f(x), \varepsilon)$. Hence, $\lim _{n \rightarrow \infty} f\left(x_{n}\right)=f(x)$.

Assume on the other hand that $f$ is sequentially continuous at $x$. If $f$ was not continuous in $x$ then there exists $\varepsilon>0$ such that for every $n \in \mathbb{N}$ there exists $x_{n} \in B\left(x, \frac{1}{n}\right)$ with $f\left(x_{n}\right) \notin B(f(x), \boldsymbol{\varepsilon})$. By construction, $\lim _{n \rightarrow \infty} x_{n}=x$. Since $f$ is sequentially continuous, $\lim _{n \rightarrow \infty} f\left(x_{n}\right)=f(x)$. But this is a contradiction to $f\left(x_{n}\right) \notin B(f(x), \boldsymbol{\varepsilon})$, and therefore $f$ is continuous.

Lemma A.25. A function $f: M_{1} \rightarrow M_{2}$ between two metric spaces is continuous if and only if preimages of open sets are open, that is, if and only if for every open set $O \subseteq M_{2}$ the preimage $f^{-1}(O)$ is open in $M_{1}$.

Proof. Let $f: M_{1} \rightarrow M_{2}$ be continuous and let $O \subseteq M_{2}$ be open. Let $x \in f^{-1}(O)$. Since $O$ is open, there exists $\varepsilon>0$ such that $B(f(x), \varepsilon) \subseteq O$. Since $f$ is continuous, there exists $\delta>0$ such that for every $y \in B(x, \delta)$ one has $f(y) \in B(f(x), \varepsilon)$. Hence, $B(x, \delta) \subseteq f^{-1}(O)$ so that $f^{-1}(O)$ is open.

On the other hand, if the preimage of every open set is open, then for every $x \in M_{1}$ and every $\varepsilon>0$ the preimage $f^{-1}(B(f(x), \varepsilon))$ is open. Clearly, $x$ belongs to this preimage, and therefore there exists $\delta>0$ such that $B(x, \delta) \subseteq f^{-1}(B(f(x), \varepsilon))$. This proves continuity.

Lemma A.26. Let $f: K \rightarrow M$ be a continuousfunction from a compact metric space $K$ into a metric space $M$. Then:
a) The image $f(K)$ is compact.
b) The function $f$ is uniformly continuous.

Proof. (a) Let $\left(O_{i}\right)_{i \in I}$ be an open covering of $f(K)$. Since $f$ is continuous, $f^{-1}\left(O_{i}\right)$ is open in $K$. Moreover, $\left(f^{-1}\left(O_{i}\right)\right)_{i \in I}$ is an open covering of $K$. Since $K$ is compact, there exists a finite subcovering: $K=\bigcup_{i \in I_{0}} f^{-1}\left(O_{i}\right)$ for some finite $I_{0} \subseteq I$. Hence, $\left(O_{i}\right)_{i \in I_{0}}$ is a finite subcovering of $f(K)$.
(b) Let $\varepsilon>0$. Since $f$ is continuous, for every $x \in K$ there exists $\delta_{x}>0$ such that for all $y \in B\left(x, \delta_{x}\right)$ one has $f(y) \in B(f(x), \varepsilon)$. By compactness, there exists a finite family $\left(x_{i}\right)_{1 \leq i \leq n} \subseteq K$ such that $K=\bigcup_{i=1}^{n} B\left(x_{i}, \delta_{x_{i}} / 2\right)$. Let $\delta=\min \left\{\delta_{x_{i}} / 2: 1 \leq i \leq n\right\}$ and let $x, y \in K$ such that $d(x, y)<\delta$. Since $x \in B\left(x_{i}, \delta_{x_{i}} / 2\right)$ for some $1 \leq i \leq n$, we find that $y \in B\left(x_{i}, \delta_{x_{i}}\right)$. By construction, $f(x), f(y) \in B\left(f\left(x_{i}\right), \varepsilon\right)$ so that the triangle inequality implies $d(f(x), f(y))<2 \varepsilon$.

Lemma A.27. Any Lipschitz continuous function $f: M_{1} \rightarrow M_{2}$ between two metric spaces is uniformly continuous.

Proof. Let $L>0$ be a Lipschitz constant for $f$ and let $\varepsilon>0$. Define $\delta:=\varepsilon / L$. Then, for every $x, y \in M$ such that $d_{1}(x, y) \leq \delta$ one has

$$
d_{2}(f(x), f(y)) \leq L d_{1}(x, y) \leq \varepsilon
$$

and therefore $f$ is uniformly continuous.

## A. 5 Completion of a metric space

Definition A.28. We say that a subset $D \subseteq M$ of a metric space $(M, d)$ is dense in $M$ if $\bar{D}=M$. Equivalently, $D$ is dense in $M$ if for every $x \in M$ there exists $\left(x_{n}\right) \subseteq D$ such that $\lim _{n \rightarrow \infty} x_{n}=x$.

Lemma A. 29 (Completion). Let $(M, d)$ be a metric space. Then there exists a complete metric space $(\hat{M}, \hat{d})$ and a continuous, injective $j: M \rightarrow \hat{M}$ such that

$$
d(x, y)=\hat{d}(j(x), j(y)), \quad x, y \in M
$$

and such that the image $j(M)$ is dense in $\hat{M}$.
Definition A.30. Let $(M, d)$ be a metric space. A complete metric space $(\hat{M}, \hat{d})$ fulfilling the properties from Lemma A. 29 is called a completion of $M$.

Proof (Proof of Lemma A.29). Let

$$
\bar{M}:=\left\{\left(x_{n}\right) \subseteq M:\left(x_{n}\right) \text { is a Cauchy sequence }\right\} .
$$

We say that two Cauchy sequences $\left(x_{n}\right),\left(y_{n}\right) \subseteq \bar{M}$ are equivalent (and we write $\left.\left(x_{n}\right) \sim\left(y_{n}\right)\right)$ if $\lim _{n \rightarrow \infty} d\left(x_{n}, y_{n}\right)=0$. Clearly, $\sim$ is an equivalence relation on $\bar{M}$.

We denote by $\left[\left(x_{n}\right)\right]$ the equivalence class in $\bar{M}$ of a Cauchy sequence $\left(x_{n}\right)$, and we let

$$
\hat{M}:=\bar{M} / \sim=\left\{\left[\left(x_{n}\right)\right]:\left(x_{n}\right) \in \bar{M}\right\}
$$

be the set of all equivalence classes. If we define

$$
\hat{d}\left(\left[\left(x_{n}\right)\right],\left[\left(y_{n}\right)\right]\right):=\lim _{n \rightarrow \infty} d\left(x_{n}, y_{n}\right),
$$

then $\hat{d}$ is well defined (the definition is independent of the choice of representatives) and it is a metric on $\hat{M}$. The fact that $\hat{d}$ is a metric and also that $(\hat{M}, \hat{d})$ is a complete metric space are left as exercises.

One also easily verifies that $j: M \rightarrow \hat{M}$ defined by $j(x)=[(x)]$ (the equivalence class of the constant sequence $(x)$ ) is continuous, injective and in fact isometric, i.e.

$$
d(x, y)=\hat{d}(j(x), j(y))
$$

for every $x, y \in M$. The proof is here complete.
Lemma A.31. Let $\left(\hat{M}_{i}, \hat{h}_{i}\right)(i=1,2)$ be two completions of a metric space $(M, d)$. Then there exists a bijection $b: \hat{M}_{1} \rightarrow \hat{M}_{2}$ such that for every $x, y \in \hat{M}_{1}$

$$
\hat{d}_{1}(x, y)=\hat{d}_{2}(b(x), b(y)) .
$$

Lemma A. 31 shows that up to isometric bijections there exists only one completion of a given metric space and it allows us to speak of the completion of a metric space.

Lemma A.32. Let $f: M_{1} \rightarrow M_{2}$ be a uniformly (!) continuous function between two metric spaces. Let $\hat{M}_{1}$ and $\hat{M}_{2}$ be the completions of $M_{1}$ and $M_{2}$, respectively. Then there exists a unique continuous extension $\hat{f}: \hat{M}_{1} \rightarrow \hat{M}_{2}$ of $f$.

Proof. Since $f$ is uniformly continuous, it maps equivalent Cauchy sequences into equivalent Cauchy sequences (equivalence of Cauchy sequences is defined as in the proof of Lemma A.29). Hence, the function $\hat{f}\left(\left[\left(x_{n}\right)\right]\right):=\left[\left(f\left(x_{n}\right)\right)\right]$ is well defined. It is easy to check that $\hat{f}$ is an extension of $f$ and that $\hat{f}$ is continuous (even uniformly continuous).

The assumption of uniform continuity in Lemma A. 32 is necessary in general. The functions $f(x)=\sin (1 / x)$ and $f(x)=1 / x$ on the open interval $(0,1)$ do not admit continuous extensions to the closed interval $[0,1]$ (which is the completion of $(0,1))$.

## Appendix B

## Banach spaces and bounded linear operators

Throughout, let $\mathbb{K} \in\{\mathbb{R}, \mathbb{C}\}$.

## B. 1 Normed spaces

Definition B.1. Let $X$ be a vector space over $\mathbb{K}$. A function $\|\cdot\|: X \rightarrow \mathbb{R}_{+}$is called a norm if for every $x, y \in X$ and every $\lambda \in \mathbb{K}$
(i) $\|x\|=0$ if and only if $x=0$,
(ii) $\|\lambda x\|=|\lambda|\|x\|$, and
(iii) $\|x+y\| \leq\|x\|+\|y\|$ (triangle inequality).

A pair $(X,\|\cdot\|)$ of a vector space $X$ and a norm $\|\cdot\|$ is called a normed space.
Often, we will speak of a normed space $X$ if it is clear which norm is given on $X$.
Example B.2. 1. (Finite dimensional spaces) Let $X=\mathbb{K}^{d}$. Then

$$
\|x\|_{p}:=\left(\sum_{i=1}^{d}\left|x_{i}\right|^{p}\right)^{1 / p}, \quad 1 \leq p<\infty
$$

and

$$
\|x\|_{\infty}:=\sup _{1 \leq i \leq d}\left|x_{i}\right|
$$

are norms on $X$.
2. (Sequence spaces) Let $1 \leq p<\infty$, and let

$$
l^{p}:=\left\{\left(x_{n}\right) \subset \mathbb{K}: \sum_{n}\left|x_{n}\right|^{p}<\infty\right\}
$$

with norm

$$
\|x\|_{p}:=\left(\sum_{n}\left|x_{n}\right|^{p}\right)^{1 / p}
$$

Then $\left(l^{p},\|\cdot\|_{p}\right)$ is a normed space.
3. (Sequence spaces) Let $X$ be one of the spaces

$$
\begin{aligned}
l^{\infty} & :=\left\{\left(x_{n}\right) \subset \mathbb{K}: \sup _{n}\left|x_{n}\right|<\infty\right\} \\
c & :=\left\{\left(x_{n}\right) \subset \mathbb{K}: \lim _{n \rightarrow \infty} x_{n} \text { exists }\right\}, \text { or } \\
c_{0} & :=\left\{\left(x_{n}\right) \subset \mathbb{K}: \lim _{n \rightarrow \infty} x_{n}=0\right\}, \text { or } \\
c_{00} & :=\left\{\left(x_{n}\right) \subset \mathbb{K}: \text { the set }\left\{n: x_{n} \neq 0\right\} \text { is finite }\right\},
\end{aligned}
$$

and let

$$
\|x\|_{\infty}:=\sup _{n}\left|x_{n}\right|
$$

Then $\left(X,\|\cdot\|_{\infty}\right)$ is a normed space.
4. (Function spaces: continuous functions) Let $C([a, b])$ be the space of all continuous, $\mathbb{K}$-valued functions on a compact interval $[a, b] \subset \mathbb{R}$. Then

$$
\|f\|_{p}:=\left(\int_{a}^{b}|f(x)|^{p} d x\right)^{1 / p}, \quad 1 \leq p<\infty
$$

and

$$
\|f\|_{\infty}:=\sup _{x \in[a, b]}|f(x)|
$$

are norms on $C([a, b])$.
5. (Function spaces: continuous functions) Let $K$ be a compact metric space and let $C(K)$ be the space of all continuous, $\mathbb{K}$-valued functions on $K$. Then

$$
\|f\|_{\infty}:=\sup _{x \in K}|f(x)|
$$

is a norm on $C(K)$.
6. (Function spaces: integrable functions) $\operatorname{Let}(\Omega, \mathscr{A}, \mu)$ be a measure space and let $X_{p}=L^{p}(\Omega)(1 \leq p \leq \infty)$. Let

$$
\|f\|_{p}:=\left(\int_{\Omega}|f|^{p} d \mu\right)^{1 / p}, \quad 1 \leq p<\infty
$$

or

$$
\|f\|_{\infty}:=\operatorname{ess} \sup |f(x)|:=\inf \left\{c \in \mathbb{R}_{+}: \mu(\{|f|>c\})=0\right\}
$$

Then $\left(X_{p},\|\cdot\|_{p}\right)$ is a normed space.
7. (Function spaces: differentiable functions) Let

$$
C^{1}([a, b]):=\{f \in C([a, b]): f \text { is continuously differentiable }\} .
$$

Then $\|\cdot\|_{\infty}$ and

$$
\|f\|_{C^{1}}:=\|f\|_{\infty}+\left\|f^{\prime}\right\|_{\infty}
$$

are norms on $C^{1}([a, b])$.
We will see more examples in the sequel.
Lemma B.3. Every normed space is a metric space for the metric

$$
d(x, y):=\|x-y\|, \quad x, y \in X
$$

By the above lemma, also every subset of a normed space becomes a metric space in a natural way. Moreover, it is natural to speak of closed or open subsets (or linear subspaces!) of normed spaces, or of closures and interiors of subsets.

Exercise B. 4 Show that in a normed space $X$, for every $x \in X$ and every $r>0$ the closed ball $\bar{B}(x, r)$ coincides with closure $\overline{B(x, r)}$ of the open ball.

Also the notion of continuity of functions between normed spaces (or between a metric space and a normed space) makes sense. The following is a first example of a continuous function.

Lemma B.5. Given a normed space, the norm is a continuous function.
This lemma is a consequence of the following lemma.
Lemma B. 6 (Triangle inequality from below). Let $X$ be a normed space. Then, for every $x, y \in X$,

$$
\|x-y\| \geq|\|x\|-\|y\|| .
$$

Proof. The triangle inequality implies

$$
\begin{aligned}
\|x\| & =\|x-y+y\| \\
& \leq\|x-y\|+\|y\|
\end{aligned}
$$

so that

$$
\|x\|-\|y\| \leq\|x-y\| .
$$

Changing the role of $x$ and $y$ implies

$$
\|y\|-\|x\| \leq\|y-x\|=\|x-y\|,
$$

and the claim follows.
A notion which can not really be defined in metric spaces but in normed spaces is the following.

Definition B.7. A subset $B$ of a normed space $X$ is called bounded if

$$
\sup \{\|x\|: x \in B\}<\infty
$$

It is easy to check that if $X$ is a normed space, and $M$ is a metric space, then the set $C(M ; X)$ of all continuous functions from $M$ into $X$ is a vector space for the obvious addition and scalar multiplication. If $M$ is in addition compact, then $f(M) \subset X$ is also compact for every such function, and hence $f(M)$ is necessarily bounded (every compact subset of a normed space is bounded!). So we can give a new example of a normed space.

Example B.8. 8. (Function spaces: vector-valued continuous functions) Let ( $X, \|$. $\|)$ be a normed space and let $K$ be a compact metric space. Let $E=C(K ; X)$ be the space of all $X$-valued continuous functions on $K$. Then

$$
\|f\|_{\infty}:=\sup _{x \in K}\|f(x)\|
$$

is a norm on $C(K ; X)$.
Also the notions of Cauchy sequences and convergent sequences make sense in normed spaces. In particular, one can speak of a complete normed space, that is, a normed space in which every Cauchy sequence converges.

Definition B.9. A complete normed space is called a Banach space.
Example B.10. The finite dimensional spaces, the sequence spaces $l^{p}(1 \leq p \leq \infty)$, $c$, and $c_{0}$, and the function spaces $\left(C([a, b]),\|\cdot\|_{\infty}\right),\left(L^{p}(\Omega),\|\cdot\|_{p}\right)$ are Banach spaces.

The spaces $\left(c_{00},\|\cdot\|_{\infty}\right),\left(C([a, b]),\|\cdot\|_{p}\right)(1 \leq p<\infty)$ are not Banach spaces. If $X$ is a Banach space, then also $\left(C(K ; X),\|\cdot\|_{\infty}\right)$ is a Banach space.

Definition B.11. We say that two norms $\|\cdot\|_{1}$ and $\|\cdot\|_{2}$ on a real or complex vector space $X$ are equivalent if there exist two constants $c, C>0$ such that for every $x \in X$

$$
c\|x\|_{1} \leq\|x\|_{2} \leq C\|x\|_{1}
$$

Lemma B.12. Let $\|\cdot\|_{1},\|\cdot\|_{2}$ be two norms on a vector space $X$ (over $\mathbb{K}$ ). The following are equivalent:
(i) The norms $\|\cdot\|_{1},\|\cdot\|_{2}$ are equivalent.
(ii) A set $O \subset X$ is open for the norm $\|\cdot\|_{1}$ if and only if it is open for the norm $\|\cdot\|_{2}$ (and similarly for closed sets).
(iii) A sequence $\left(x_{n}\right) \subset X$ converges to 0 for the norm $\|\cdot\|_{1}$ if and only if it converges to 0 for the norm $\|\cdot\|_{2}$.

In other words, if two norms $\|\cdot\|_{1},\|\cdot\|_{2}$ on a vector space $X$ are equivalent, then the open sets, the closed sets and the null sequences are the same. We also say that the two norms define the same topology. In particular, if $X$ is a Banach space for one norm then it is also a Banach space for the other (equivalent) norm.

Exercise B. 13 The norms $\|\cdot\|_{\infty}$ and $\|\cdot\|_{p}$ are not equivalent on $C([0,1])$.

Theorem B.14. Any two norms on a finite dimensional real or complex vector space are equivalent.

Proof. We may without loss of generality consider $\mathbb{K}^{d}$. Let $\|\cdot\|$ be a norm on $\mathbb{K}^{d}$ and let $\left(e_{i}\right)_{1 \leq i \leq d}$ be the canonical basis of $\mathbb{K}^{d}$. For every $x \in \mathbb{K}^{d}$

$$
\begin{aligned}
\|x\| & =\left\|\sum_{i=1}^{d} x_{i} e_{i}\right\| \\
& \leq \sum_{i=1}^{d}\left|x_{i}\right|\left\|e_{i}\right\| \\
& \leq C\|x\|_{1},
\end{aligned}
$$

where $C:=\sup _{1 \leq i \leq d}\left\|e_{i}\right\|<\infty$ and $\|\cdot\|_{1}$ is the norm from Example B.2.1. By the triangle inequality from below, for every $x, y \in \mathbb{K}^{d}$,

$$
\mid\|x\|-\|y\|\|\leq\| x-y\|\leq C\| x-y \|_{1}
$$

Hence, the norm $\|\cdot\|:\left(\mathbb{K}^{d},\|\cdot\|_{1}\right) \rightarrow \mathbb{R}_{+}$is continuous (on $\mathbb{K}^{d}$ equipped with the norm $\|\cdot\|_{1}$ ). If $S:=\left\{x \in \mathbb{K}^{d}:\|x\|_{1}=1\right\}$ denotes the unit sphere for the norm $\|\cdot\|_{1}$, then $S$ is compact. As a consequence

$$
c:=\inf \{\|x\|: x \in S\}>0
$$

since the infimum is attained by the continuity of $\|\cdot\|$. This implies

$$
c\|x\|_{1} \leq\|x\| \quad \text { for every } x \in \mathbb{K}^{d}
$$

We have proved that every norm on $\mathbb{K}^{d}$ is equivalent to the norm $\|\cdot\|_{1}$. Hence, any two norms on $\mathbb{K}^{d}$ are equivalent.

Corollary B.15. Any finite dimensional normed space is complete. Any finite dimensional subspace of a normed space is closed.

Proof. The space $\left(\mathbb{K}^{d},\|\cdot\|_{1}\right)$ is complete (exercise!). If $\|\cdot\|$ is a second norm on $\mathbb{K}^{d}$ and if $\left(x_{n}\right)$ is a Cauchy sequence for that norm, then it is also a Cauchy sequence in $\left(\mathbb{K}^{d},\|\cdot\|_{1}\right)$ (use that the norms $\|\cdot\|_{1}$ and $\|\cdot\|$ are equivalent), and therefore convergent in $\left(\mathbb{K}^{d},\|\cdot\|_{1}\right)$. By equivalence of norms again, the sequence $\left(x_{n}\right)$ is also convergent in $\left(\mathbb{K}^{d},\|\cdot\|\right)$, and therefore $\left(\mathbb{K}^{d},\|\cdot\|\right)$ is complete.

Let $Y$ be a finite dimensional subspace of a normed space $X$, and let $\left(x_{n}\right) \subset Y$ be a convergent sequence with $x=\lim _{n \rightarrow \infty} x_{n} \in X$. Since $\left(x_{n}\right)$ is also a Cauchy sequence, and since $Y$ is complete, we find (by uniqueness of the limit) that $x \in Y$, and therefore $Y$ is closed (Lemma A.15).

Definition B.16. Let $\left(x_{n}\right)$ be a sequence in a normed space $X$. We say that the series $\sum_{n} x_{n}$ is convergent if the sequence $\left(\sum_{j \leq n} x_{j}\right)$ of partial sums is convergent. We say that the series $\sum_{n} x_{n}$ is absolutely convergent if $\sum_{n}\left\|x_{n}\right\|<\infty$.

Lemma B.17. Let $\left(x_{n}\right)$ be a sequence in a normed space $X$. If the series $\sum_{n} x_{n}$ is convergent, then necessarily $\lim _{n \rightarrow \infty} x_{n}=0$.

Note that in a normed space not every absolutely convergent series is convergent. In fact, the following is true.

Lemma B.18. A normed space $X$ is a Banach space if and only if every absolutely convergent series converges.

Proof. Assume that $X$ is a Banach space, and let $\sum_{n} x_{n}$ be absolutely convergent. It follows easily from the triangle inequality that the corresponding sequence of partial sums is a Cauchy sequence, and since $X$ is complete, the series $\sum_{n} x_{n}$ is convergent.

On the other hand, assume that every absolutely convergent series is convergent. Let $\left(x_{n}\right)_{n \geq 1} \subset X$ be a Cauchy sequence. From this Cauchy sequence, one can extract a subsequence $\left(x_{n_{k}}\right)_{k \geq 1}$ such that $\left\|x_{n_{k+1}}-x_{n_{k}}\right\| \leq 2^{-k}, k \geq 1$. Let $y_{0}=x_{n_{1}}$ and $y_{k}=x_{n_{k+1}}-x_{n_{k}}, k \geq 1$. Then the series $\sum_{k \geq 0} y_{k}$ is absolutely convergent. By assumption, it is also convergent. But by construction, $\left(\sum_{l=0}^{k} y_{l}\right)=\left(x_{n_{k}}\right)$, so that $\left(x_{n_{k}}\right)$ is convergent. Hence, we have extracted a subsequence of the Cauchy sequence $\left(x_{n}\right)$ which converges. As a consequence, $\left(x_{n}\right)$ is convergent, and since $\left(x_{n}\right)$ was an arbitrary Cauchy sequence, $X$ is complete.

Lemma B. 19 (Riesz). Let $X$ be a normed space and let $Y \subset X$ be a closed linear subspace. If $Y \neq X$, then for every $\delta>0$ there exists $x \in X \backslash Y$ such that $\|x\|=1$ and

$$
\operatorname{dist}(x, Y)=\inf \{\|x-y\|: y \in Y\} \geq 1-\delta
$$

Proof. Let $z \in X \backslash Y$. Since $Y$ is closed,

$$
d:=\operatorname{dist}(z, Y)>0
$$

Let $\delta>0$. By definition of the infimum, there exists $y \in Y$ such that

$$
\|z-y\| \leq \frac{d}{1-\delta}
$$

Let $x:=\frac{z-y}{\|z-y\|}$. Then $x \in X \backslash Y,\|x\|=1$, and for every $u \in Y$

$$
\begin{aligned}
\|x-u\| & =\|z-y\|^{-1}\|z-(y+\|z-y\| u)\| \\
& \geq\|z-y\|^{-1} d \geq 1-\delta,
\end{aligned}
$$

since $(y+\|z-y\| u) \in Y$.
Theorem B.20. A normed space is finite dimensional if and only if every closed bounded set is compact.

Proof. If the normed space is finite dimensional, then every closed bounded set is compact by the Theorem of Heine-Borel. Note that by Theorem B. 14 it is not
important which norm on the finite dimensional space is considered. By Lemma B. 12 , the closed and bounded sets do not change.

On the other hand, if the normed space is infinite dimensional, then, by the Lemma of Riesz, one can construct inductively a sequence $\left(x_{n}\right) \subset X$ such that $\left\|x_{n}\right\|=1$ and $\operatorname{dist}\left(x_{n+1}, X_{n}\right) \geq \frac{1}{2}$ for every $n \in \mathbb{N}$, where $X_{n}=\operatorname{span}\left\{x_{i}: 1 \leq i \leq n\right\}$ (note that $X_{n}$ is closed by Corollary B.15). By construction, $\left(x_{n}\right)$ belongs to the closed unit ball, but it can not have a convergent subsequence (even not a Cauchy subsequence). Hence, the closed unit ball is not compact. We state this result separately.

Theorem B.21. In an infinite dimensional Banach space the closed unit ball is not compact.

Lemma B. 22 (Completion of a normed space). For every normed space $X$ there exists a Banach space $\hat{X}$ and a linear injective $j: X \rightarrow \hat{X}$ such that $\|j(x)\|=\|x\|$ $(x \in X)$ and $j(X)$ is dense in $\hat{X}$. Up to isometry, the Banach space $\hat{X}$ is unique (up to isomorphism). It is called the completion of $X$.

Proof. It suffices to repeat the proof of Lemma A. 29 and to note that the completion $\hat{X}$ of $X$ (considered as a metric space) carries in a natural way a linear structure: addition of - equivalence classes of - Cauchy sequences is their componentwise addition, and also multiplication of - an equivalence class - of a Cauchy sequence and a scalar is done componentwise. Moreover, for every $\left[\left(x_{n}\right)\right]$, one defines the norm

$$
\left\|\left[\left(x_{n}\right)\right]\right\|:=\lim _{n \rightarrow \infty}\left\|x_{n}\right\| .
$$

Uniqueness of $\hat{X}$ follows from Lemma A. 31 .

## B. 2 Product spaces and quotient spaces

Lemma B. 23 (Product spaces). Let $\left(X_{i}\right)_{i \in I}$ be a finite (!) family of normed spaces, and let $\mathscr{X}:=\bigotimes_{i \in I} X_{i}$ be the cartesian product. Then

$$
\|x\|_{p}:=\left(\sum_{i \in I}\left\|x_{i}\right\|_{X_{i}}^{p}\right)^{1 / p} \quad(1 \leq p<\infty)
$$

and

$$
\|x\|_{\infty}:=\sup _{i \in I}\left\|x_{i}\right\|_{X_{i}}
$$

define equivalent norms on $\mathscr{X}$. In particular, the cartesian product is a normed space.

Proof. The easy proof is left to the reader.

Lemma B.24. Let $\left(X_{i}\right)_{i \in I}$ be a finite family of normed spaces, and let $\mathscr{X}:=\bigotimes_{i \in I} X_{i}$ be the cartesian product equipped with one of the equivalent norms $\|\cdot\|_{p}$ from Lemma B.23. Then a sequence $\left(x^{n}\right)=\left(\left(x_{i}^{n}\right)_{i}\right) \subset \mathscr{X}$ converges (is a Cauchy sequence) if and only if $\left(x_{i}^{n}\right) \subset X_{i}$ is convergent (is a Cauchy sequence) for every $i \in I$.

As a consequence, $\mathscr{X}$ is a Banach space if and only if all the $X_{i}$ are Banach spaces.

Proposition B. 25 (Quotient space). Let $X$ be a vector space (!) over $\mathbb{K}$, and let $Y \subset X$ be a linear subspace. Define, for every $x \in X$, the affine subspace

$$
x+Y:=\{x+y: y \in Y\}
$$

and define the quotient space or factor space

$$
X / Y:=\{x+Y: x \in X\} .
$$

Then $X / Y$ is a vector space for the addition

$$
(x+Y)+(z+Y):=(x+z+Y)
$$

and the scalar multiplication

$$
\lambda(x+Y):=(\lambda x+Y) .
$$

The neutral element is $Y$.
For the definition of quotient spaces, it is not important that we consider real or complex vector spaces.

Examples of quotient spaces are already known. In fact, $L^{p}$ is such an example. Usually, one defines

$$
\mathscr{L}^{p}(\Omega, \mathscr{A}, \mu)
$$

to be the space of all mesurable functions $f: \Omega \rightarrow \mathbb{K}$ such that $\int_{\Omega}|f|^{p} d \mu<\infty$. Moreover,

$$
N:=\left\{f \in \mathscr{L}^{p}(\Omega, \mathscr{A}, \mu): \int_{\Omega}|f|^{p}=0\right\} .
$$

Note that $N$ is a linear subspace of $\mathscr{L}^{p}(\Omega, \mathscr{A}, \mu)$, and that $N$ is the space of all functions $f \in \mathscr{L}^{p}$ which vanish almost everywhere. Then

$$
L^{p}(\Omega, \mathscr{A}, \mu):=\mathscr{L}^{p}(\Omega, \mathscr{A}, \mu) / N
$$

Proposition B.26. Let $X$ be a normed space and let $Y \subset X$ be a linear subspace . Then

$$
\|x+Y\|:=\inf \{\|x-y\|: y \in Y\}
$$

defines a norm on $X / Y$ if and only if $Y$ is closed in $X$. If $X$ is a Banach space and $Y \subset X$ closed, then $X / Y$ is also a Banach space.
Proof. We have to check that $\|\cdot\|$ satisfies all properties of a norm. Recall that $0_{X / Y}=Y$, and that for all $x \in X$

$$
\begin{aligned}
& \|x+Y\|=0 \\
\Leftrightarrow & \inf \{\|x-y\|: y \in Y\}=0 \\
\Leftrightarrow & \exists\left(y_{n}\right) \subset Y: \lim _{n \rightarrow \infty} y_{n}=x \\
\Leftrightarrow & (\Rightarrow \text { if } Y \text { closed }): x \in Y \\
\Leftrightarrow & x+Y=Y .
\end{aligned}
$$

Second, for every $x \in X$ and every $\lambda \in \mathbb{K} \backslash\{0\}$,

$$
\begin{aligned}
\|\lambda(x+Y)\| & =\|\lambda x+Y\| \\
& =\inf \{\|\lambda x-y\|: y \in Y\} \\
& =\inf \{\|\lambda(x-y)\|: y \in Y\} \\
& =|\lambda| \inf \{\|x-y\|: y \in Y\} \\
& =|\lambda|\|x+Y\| .
\end{aligned}
$$

Third, for every $x, z \in X$,

$$
\begin{aligned}
\|(x+Y)+(z+Y)\| & =\|(x+z)+Y\| \\
& =\inf \{\|x+z-y\|: y \in Y\} \\
& =\inf \left\{\left\|x+z-y_{1}-y_{2}\right\|: y_{1}, y_{2} \in Y\right\} \\
& \leq \inf \left\{\left\|x-y_{1}\right\|+\left\|z-y_{2}\right\|: y_{1}, y_{2} \in Y\right\} \\
& \leq \inf \{\|x-y\|: y \in Y\}+\inf \{\|z-y\|: y \in Y\} \\
& =\|x+Y\|+\|z+Y\| .
\end{aligned}
$$

Hence, $X / Y$ is a normed space if $Y$ is closed.
Assume next that $X$ is a Banach space. Let $\left(x_{n}\right) \subset X$ be such that the series $\sum_{n \geq 1} x_{n}+Y$ converges absolutely, i.e. $\sum_{n \geq 1}\left\|x_{n}+Y\right\|<\infty$. By definition of the norm in $X / Y$, we find $\left(y_{n}\right) \subset Y$ such that $\left\|x_{n}-y_{n}\right\| \leq\left\|x_{n}+Y\right\|+2^{-n}$. Replacing $\left(x_{n}\right)$ by $\left(\hat{x}_{n}\right)=\left(x_{n}-y_{n}\right)$, we find that $x_{n}+Y=\hat{x}_{n}+Y$ and that the series $\sum_{n \geq 0} \hat{x}_{n}$ is absolutely convergent. Since $X$ is complete, by Lemma B.18, the limit $\sum_{n \geq 1} \hat{x}_{n}=x \in X$ exists. As a consequence,

$$
\begin{aligned}
\left\|(x+Y)-\sum_{k=1}^{n}\left(\hat{x}_{k}+Y\right)\right\| & =\left\|\left(x-\sum_{k=1}^{n} \hat{x}_{k}\right)+Y\right\| \\
& \leq\left\|x-\sum_{k=1}^{n} \hat{x}_{k}\right\| \quad \rightarrow \quad 0
\end{aligned}
$$

i.e. the series $\sum_{n \geq 1} x_{n}+Y$ converges. By Lemma B.18, $X / Y$ is complete.

## B. 3 Bounded linear operators

In the following a linear mapping between two normed spaces $X$ and $Y$ will also be called a linear operator or just operator. If $Y=\mathbb{K}$, then we call linear operators also linear functionals. If $T: X \rightarrow Y$ is a linear operator between two normed spaces, then we denote by

$$
\operatorname{ker} T:=\{x \in X: T x=0\}
$$

its kernel or null space, and by

$$
\operatorname{ran} T:=\{T x: x \in X\}
$$

its range or image. Observe that we simply write $T x$ instead of $T(x)$, meaning that $T$ is applied to $x \in X$. The identity operator $X \rightarrow X, x \mapsto x$ is denoted by $I$.

Lemma B.27. Let $T: X \rightarrow Y$ be a linear operator between two normed spaces $X$ and $Y$. Then the following are equivalent
(i) $T$ is continuous.
(ii) $T$ is continuous at 0 .
(iii) $T B$ is bounded in $Y$, where $B=B(0,1)$ denotes the unit ball in $X$.
(iv) There exists a constant $C \geq 0$ such that for every $x \in X$

$$
\|T x\| \leq C\|x\|
$$

Proof. The implication (i) $\Rightarrow$ (ii) is trivial.
(ii) $\Rightarrow$ (iii). If $T$ is continuous at 0 , then there exists some $\delta>0$ such that for every $x \in B(0, \delta)$ one has $T x \in B(0,1)$ (so the $\varepsilon$ from the $\varepsilon$ - $\delta$ definition of continuity is chosen to be 1 here). By linearity, for every $x \in B=B(0,1)$

$$
\|T x\|=\frac{1}{\delta}\|T(\delta x)\| \leq \frac{1}{\delta}
$$

and this means that $T B$ is bounded.
(iii) $\Rightarrow$ (iv). The set $T B$ being bounded in $Y$ means that there exists some constant $C \geq 0$ such that for every $x \in B$ one has $\|T x\| \leq C$. By linearity, for every $x \in X \backslash\{0\}$,

$$
\|T x\|=\left\|T \frac{x}{\|x\|}\right\|\|x\| \leq C\|x\|
$$

(iv) $\Rightarrow$ (i). Let $x \in X$, and assume that $\lim _{n \rightarrow \infty} x_{n}=x$. Then

$$
\left\|T x_{n}-T x\right\|=\left\|T\left(x_{n}-x\right)\right\| \leq C\left\|x_{n}-x\right\| \rightarrow 0 \quad \text { as } n \rightarrow \infty
$$

so that $\lim _{n \rightarrow \infty} T x_{n}=T x$.
Definition B.28. We call a continuous linear operator $T: X \rightarrow Y$ between two normed spaces $X$ and $Y$ also a bounded operator (since it maps the unit ball of
$X$ to a bounded subset of $Y$ ). The set of all bounded linear operators is denoted by $\mathscr{L}(X, Y)$. Special cases: If $X=Y$, then we write $\mathscr{L}(X, X)=: \mathscr{L}(X)$. If $Y=\mathbb{K}$, then we write $\mathscr{L}(X, \mathbb{K})=: X^{\prime}$.

Lemma B.29. The set $\mathscr{L}(X, Y)$ is a vector space and

$$
\begin{align*}
\|T\| & :=\inf \{C \geq 0:\|T x\| \leq C\|x\| \text { for all } x \in X\}  \tag{B.1}\\
& =\sup \{\|T x\|:\|x\| \leq 1\} \\
& =\sup \{\|T x\|:\|x\|=1\}
\end{align*}
$$

is a norm on $\mathscr{L}(X, Y)$.
Proof. We first show that the three quantities on the right-hand side of (B.1) are equal. In fact, the equality

$$
\sup \{\|T x\|:\|x\| \leq 1\}=\sup \{\|T x\|:\|x\|=1\}
$$

is easy to check so that it remains only to show that

$$
A:=\inf \{C \geq 0:\|T x\| \leq C\|x\| \text { for all } x \in X\}=\sup \{\|T x\|:\|x\|=1\}=: B
$$

If $C>A$, then for every $x \in X \backslash\{0\},\|T x\| \leq C\|x\|$ or $\left\|T \frac{x}{\|x\|}\right\| \leq C$. Hence, $C \geq B$ which implies that $A \geq B$. If $C>B$, then for every $x \in X \backslash\{0\},\left\|T \frac{x}{\|x\|}\right\| \leq C$, and therefore $\|T x\| \leq C\|x\|$. Hence, $C \geq A$ which implies that $A \leq B$.

Now we check that $\|\cdot\|$ is a norm on $\mathscr{L}(X, Y)$. First, for every $T \in \mathscr{L}(X, Y)$,

$$
\begin{aligned}
\|T\|=0 & \Leftrightarrow \sup \{\|T x\|:\|x\| \leq 1\}=0 \\
& \Leftrightarrow \forall x \in X,\|x\| \leq 1:\|T x\|=0 \\
& \Leftrightarrow(\|\cdot\| \text { is a norm on } Y) \forall x \in X,\|x\| \leq 1: T x=0 \\
& \Leftrightarrow(\Rightarrow \text { linearity of } T) \forall x \in X: T x=0 \\
& \Leftrightarrow T=0 .
\end{aligned}
$$

Second, for every $T \in \mathscr{L}(X, Y)$ and every $\lambda \in \mathbb{K}$

$$
\begin{aligned}
\|\lambda T\| & =\sup \{\|(\lambda T) x\|:\|x\| \leq 1\} \\
& =\sup \{|\lambda|\|T x\|:\|x\| \leq 1\} \\
& =|\lambda|\|T\|
\end{aligned}
$$

Finally, for every $T, S \in \mathscr{L}(X, Y)$,

$$
\begin{aligned}
\|T+S\| & =\sup \{\|(T+S) x\|:\|x\| \leq 1\} \\
& \leq \sup \{\|T x\|+\|S x\|:\|x\| \leq 1\} \\
& \leq\|T\|+\|S\|
\end{aligned}
$$

The proof is complete.

Remark B.30. (a) Note that the infimum on the right-hand side of (B.1) in Lemma B. 29 is always attained. Thus, for every operator $T \in \mathscr{L}(X, Y)$ and every $x \in X$,

$$
\|T x\| \leq\|T\|\|x\|
$$

This inequality shall be frequently used in the sequel! Note that on the other hand the suprema on the right-hand side of (B.1) are not always attained. (b) From Lemma B. 29 we can learn how to show that some operator $T: X \rightarrow Y$ is bounded and how to calculate the norm $\|T\|$. Usually (in most cases), one should prove in the first step some inequality of the form

$$
\|T x\| \leq C\|x\|, \quad x \in X
$$

because this inequality shows on the one hand that $T$ is bounded, and on the other hand it shows the estimate $\|T\| \leq C$. In the second step one should prove that the estimate $C$ was optimal by finding some $x \in X$ of norm $\|x\|=1$ such that $\|T x\|=C$, or by finding some sequence $\left(x_{n}\right) \subset X$ of norms $\left\|x_{n}\right\| \leq 1$ such that $\lim _{n \rightarrow \infty}\left\|T x_{n}\right\|=$ $C$, because this shows that $\|T\|=C$. Of course, the second step only works if one has not lost anything in the estimate of the first step. There are in fact many examples of bounded operators for which it is difficult to estimate their norm.
Example B.31. 1. (Shift-operator). On $l^{p}(\mathbb{N})$ consider the left-shift operator

$$
L x=L\left(x_{n}\right)=\left(x_{n+1}\right) .
$$

Then

$$
\left\|L\left(x_{n}\right)\right\|_{p}=\left(\sum_{n}\left|x_{n+1}\right|^{p}\right)^{1 / p} \leq\left(\sum_{n}\left|x_{n}\right|^{p}\right)^{1 / p}
$$

so that $L$ is bounded and $\|L\| \leq 1$. On the other hand, for $x=(0,1,0,0, \ldots)$ one computes that $\|x\|_{p}=1$ and $\|L x\|_{p}=\|(1,0,0, \ldots)\|_{p}=1$, and one concludes that $\|L\|=1$.
2. (Shift-operator). Similarly, one shows that the right-shift operator $R$ on $l^{p}(\mathbb{N})$ defined by

$$
R x=R\left(x_{n}\right)=\left(0, x_{0}, x_{1}, \ldots\right)
$$

is bounded and $\|R\|=1$. Note that actually $\|R x\|_{p}=\|x\|_{p}$ for every $x \in l^{p}$.
3. (Multiplication operator). Let $m \in l^{\infty}$ and consider on $l^{p}$ the multiplication operator

$$
M x=M\left(x_{n}\right)=\left(m_{n} x_{n}\right)
$$

4. (Functionals on $C$ ). Consider the linear functional $\varphi: C([0,1]) \rightarrow \mathbb{K}$ defined by

$$
\varphi(f):=\int_{0}^{\frac{1}{2}} f(x) d x
$$

Then

$$
|\varphi(f)| \leq \int_{0}^{\frac{1}{2}}|f(x)| d x \leq \frac{1}{2}\|f\|_{\infty}
$$

so that $\varphi$ is bounded and $\|\varphi\| \leq \frac{1}{2}$. On the other hand, for the constant function $f=1$ one has $\|f\|_{\infty}=1$ and $|\varphi(f)|=\frac{1}{2}$, so that $\|\varphi\|=\frac{1}{2}$.
Lemma B.32. Let $X, Y, Z$ be three Banach spaces, and let $T \in \mathscr{L}(X, Y)$ and $S \in$ $\mathscr{L}(Y, Z)$. Then $S T \in \mathscr{L}(X, Z)$ and

$$
\|S T\| \leq\|S\|\|T\| .
$$

Proof. The boundedness of $S T$ is clear since compositions of continuous functions are again continuous. To obtain the bound on $S T$, we calculate

$$
\begin{aligned}
\|S T\| & =\sup _{\|x\| \leq 1}\|S T x\| \\
& \leq \sup _{\|x\| \leq 1}\|S\|\|T x\| \\
& \leq\|S\|\|T\| .
\end{aligned}
$$

Lemma B.33. If $Y$ is a Banach space then $\mathscr{L}(X, Y)$ is a Banach space.
Proof. Assume that $Y$ is a Banach space and let $\left(T_{n}\right)$ be a Cauchy sequence in $\mathscr{L}(X, Y)$. By the estimate

$$
\left\|T_{n} x-T_{m} x\right\|=\left\|\left(T_{n}-T_{m}\right) x\right\| \leq\left\|T_{n}-T_{m}\right\|\|x\|,
$$

the sequence $\left(T_{n} x\right)$ is a Cauchy sequence in $Y$ for every $x \in X$. Since $Y$ is complete, the limit $\lim _{n \rightarrow \infty} T_{n} x$ exists for every $x \in X$. Define $T x:=\lim _{n \rightarrow \infty} T_{n} x$. Clearly, $T$ : $X \rightarrow Y$ is linear. Moreover, since any Cauchy sequence is bounded, we find that

$$
\|T x\| \leq \sup _{n}\left\|T_{n} x\right\| \leq C\|x\|
$$

for some constant $C \geq 0$, i.e. $T$ is bounded. Moreover, for every $n \in \mathbb{N}$ we have the estimate

$$
\begin{aligned}
\left\|T-T_{n}\right\| & =\sup _{\|x\| \leq 1}\left\|T x-T_{n} x\right\| \\
& \leq \sup _{\|x\| \leq 1 m \geq n} \sup _{n}\left\|T_{m} x-T_{n} x\right\| \\
& \leq \sup _{m \geq n}\left\|T_{m}-T_{n}\right\| .
\end{aligned}
$$

Since that right-hand side of this inequality becomes arbitrarily small for large $n$, we see that $\lim _{n \rightarrow \infty} T_{n}=T$ exists, and so we have proved that $\mathscr{L}(X, Y)$ is a Banach space.
Remark B.34. The converse of the statement in Lemma B. 33 is also true, i.e if $\mathscr{L}(X, Y)$ is a Banach space then necessarily $Y$ is a Banach space. For the proof, however, one has to know that there are nontrivial operators in $\mathscr{L}(X, Y)$ as soon as $Y$ is nontrivial (i.e. $Y \neq\{0\}$ ). For this, we need the Theorem of Hahn-Banach and its consequences discussed in Chapter E.

Corollary B.35. The space $X^{\prime}=\mathscr{L}(X, \mathbb{K})$ of all bounded linear functionals on $X$ is always a Banach space.
Definition B.36. Let $X, Y$ be two normed spaces.
a) We call $T \in \mathscr{L}(X, Y)$ an isomorphism if $T$ is bijective and $T^{-1} \in \mathscr{L}(Y, X)$.
b) We call $T \in \mathscr{L}(X, Y)$ an isometry if $\|T x\|=\|x\|$ for every $x \in X$.
c) We say that $X$ and $Y$ are isomorphic (and we write $X \cong Y$ ) if there exists an isomorphism $T \in \mathscr{L}(X, Y)$.
d) We say that $X$ and $Y$ are isometrically isomorphic if there exists an isometric isomorphism $T \in \mathscr{L}(X, Y)$.

Remark B.37. 1. Two norms $\|\cdot\|_{1},\|\cdot\|_{2}$ on a $\mathbb{K}$ vector space $X$ are equivalent if and only if the identity operator $I:\left(X,\|\cdot\|_{1}\right) \rightarrow\left(X,\|\cdot\|_{2}\right)$ is an isomorphism.
2. Saying that two normed spaces $X$ and $Y$ are isomorphic means that they are not only 'equal' as vector spaces (in the sense that we find a bijective linear operator) but also as normed spaces (i.e. the bijection is continuous as well as its inverse).
3. If $T \in \mathscr{L}(X, Y)$ and $S \in \mathscr{L}(Y, Z)$ are isomorphisms, then $S T \in \mathscr{L}(X, Z)$ is an isomorphism and $(S T)^{-1}=T^{-1} S^{-1}$.
4. Every isometry $T \in \mathscr{L}(X, Y)$ is clearly injective. If it is also surjective, then $T$ is an isometric isomorphism, i.e. the inverse $T^{-1}$ is also bounded (even isometric).
5. Clearly, if $T \in \mathscr{L}(X, Y)$ is isometric, then it is an isometric isomorphism from $X$ onto $\operatorname{ran} T$, and we may say that $X$ is isometrically embedded into $Y$ (via $T$ ).

Example B.38. The right-shift operator from Example B. 31 (2) is isometric, but not surjective. In particular, $l^{p}$ is isometrically isomorphic to a proper subspace of $l^{p}$.
Exercise B. 39 Show that the spaces $\left(c,\|\cdot\|_{\infty}\right)$ of all convergent sequences and $\left(c_{0},\|\cdot\|_{\infty}\right)$ of all null sequences are isomorphic.
Exercise B. 40 Show that $\left(c_{0},\|\cdot\|_{\infty}\right)$ is (isometrically) isomorphic to a linear subspace of $\left(C([0,1]),\|\cdot\|_{\infty}\right)$, i.e. find an isometry $T: c_{0} \rightarrow C([0,1])$.
Lemma B. 41 (Neumann series). Let $X$ be a Banach space and let $T \in \mathscr{L}(X)$ be such that $\|T\|<1$. Then $I-T$ is boundedly invertible, i.e. it is an isomorphism. Moreover, $(I-T)^{-1}=\sum_{n \geq 0} T^{n}$.
Proof. Since $X$ is a Banach space, $\mathscr{L}(X)$ is also a Banach space by Lemma B.33. By assumption on $\|T\|$, the series $\sum_{n \geq 0} T^{n}$ is absolutely convergent, and hence, by Lemma B.18, it is convergent to some element $S \in \mathscr{L}(X)$. Moreover,

$$
(I-T) S=\lim _{n \rightarrow \infty}(I-T) \sum_{k=0}^{n} T^{k}=\lim _{n \rightarrow \infty}\left(I-T^{k+1}\right)=I
$$

and similarly, $S(I-T)=I$.
Corollary B.42. Let $X$ and $Y$ be two Banach spaces. Then the set $\mathscr{I}(X, Y)$ of all isomorphisms in $\mathscr{L}(X, Y)$ is open, and the mapping $T \mapsto T^{-1}$ is continuous from $\mathscr{I}(X, Y)$ onto $\mathscr{I}(Y, X)$.

Proof. Let $\mathscr{I} \subset \mathscr{L}(X, Y)$ be the set of all isomorphisms, and assume that $\mathscr{I}$ is not empty (if it is empty, then it is also open). Let $T \in \mathscr{I}$. Then for every $S \in B\left(T, \frac{1}{\left\|T^{-1}\right\|}\right)$ we have

$$
S=T+S-T=T\left(I+T^{-1}(S-T)\right),
$$

and since $\left\|T^{-1}(S-T)\right\| \leq\left\|T^{-1}\right\|\|S-T\|<1$, the operator $I+T^{-1}(S-T) \in \mathscr{L}(X)$ is an isomorphism by Lemma B.41. As a composition of two isomorphisms, $S \in \mathscr{I}$, and hence $\mathscr{I}$ is open. The continuity is also a direct consequence of the above representation of $S$ (and thus of its inverse), using the Neumann series.

## B. 4 The Arzela-Ascoli theorem

It is a consequence of Riesz' Lemma (Lemma B.19) that the unit ball in an infinite dimensional Banach space is not compact; see also Theorem B.21. But compact sets play an important role in many theorems from analysis, in particular when one wants to prove the existence of some fixed point, the existence of a solution to an algebraic equation, the existence of a solution of a differential equation, the existence of a solution of a partial differential equation etc. It is therefore important to identify the compact sets in Banach spaces, in particular in the classical Banach spaces. The Arzela-Ascoli theorem characterizes the compact subsets of $C(K ; X)$, where $(K, d)$ is a compact metric space and $X$ is a Banach space.

We say that a subset $B \subseteq C(K ; X)$ is equicontinuous at some point $x \in K$ if for every $\varepsilon>0$ there exists $\delta>0$ such that for every $y \in K$ and every $f \in B$ the implication

$$
d(x, y)<\delta \quad \Rightarrow \quad\|f(x)-f(y)\|<\varepsilon
$$

holds.
Theorem B. 43 (Arzela-Ascoli). Let $(K, d)$ be a compact metric space, $X$ be a Banach space and consider the Banach space $C(K ; X)$ of all continuous functions $K \rightarrow X$ equipped with the supremum norm $\|f\|_{\infty}=\sup _{x \in K}\|f(x)\|$. For a subset $B \subseteq C(K ; X)$, the following assertions are equivalent:
(i) The set B is compact.
(ii) The set $B$ is closed, equicontinuous at every $x \in K$ and pointwise compact in the sense that for every $x \in K$ the set $B_{x}=\{f(x): f \in B\}$ is compact.

We point out that, by the Heine-Borel theorem, the condition of pointwise compactness of $B$ can be replaced by mere pointwise boundedness or boundedness as soon as the space $X$ is finite dimensional.

Corollary B. 44 (Arzela-Ascoli). Let $(K, d)$ be a compact metric space, and consider the Banach space $C\left(K ; \mathbb{R}^{d}\right)$ of all continuous functions $K \rightarrow \mathbb{R}^{d}$ equipped with the supremum norm $\|f\|_{\infty}=\sup _{x \in K}\|f(x)\|$. For a subset $B \subseteq C\left(K ; \mathbb{R}^{d}\right)$, the following assertions are equivalent:
(i) The set B is compact.
(ii) The set $B$ is closed, equicontinuous at every $x \in K$ and pointwise bounded in the sense that for every $x \in K$ the set $B_{x}=\{f(x): f \in B\}$ is bounded.
Proof (Proof of Theorem B.43). The proof of the Arzela-Ascoli theorem is a nice application of Cantor's diagonal sequence argument which we see here for the first time, but which we will see again below when we prove that every bounded sequence in a reflexive Banach space admits a weakly convergent subsequence. Given a sequence, Cantor's argument allows us to construct a subsequence which satisfies a countable number of properties. It is instructive to learn the idea of Cantor's argument since it can be help in various situations.

We first assume that $B \subseteq C(K ; X)$ is compact. Any compact subset of a Banach space is closed and bounded, and therefore $B$ is closed and bounded, too. For every $x \in K$, the point evaluation $C(K ; X) \rightarrow X, f \mapsto f(x)$ is linear and continuous. Since continuous images of compact sets are compact, the image of $B$ under the point evaluation, that is the set $B_{x}=\{f(x): f \in B\}$, is compact.

We show that $B$ is equicontinuous at every $x$. Assume that this was not the case. Then there exist $x \in K$ and $\varepsilon>0$ such that for every $n \geq 1$ there exist $y_{n} \in K$ and $f_{n} \in B$ such that $d\left(x, y_{n}\right)<\frac{1}{n}$ and $\left\|f_{n}(x)-f_{n}\left(y_{n}\right)\right\| \geq \varepsilon$. Since $B$ is compact, there exists a subsequence of $\left(f_{n}\right)$ (which we denote for simplicity again by $\left(f_{n}\right)$ ) such that $\lim _{n \rightarrow \infty} f_{n}=f$ in $C(K ; X)$. Then, by the triangle inequality from below,

$$
\begin{aligned}
\liminf _{n \rightarrow \infty}\left\|f(x)-f\left(y_{n}\right)\right\| & =\liminf _{n \rightarrow \infty}\left\|f(x)-f_{n}(x)+f_{n}(x)-f_{n}\left(y_{n}\right)+f_{n}\left(y_{n}\right)-f\left(y_{n}\right)\right\| \\
& \geq \liminf _{n \rightarrow \infty}\left(\left\|f_{n}(x)-f_{n}\left(y_{n}\right)\right\|-2\left\|f-f_{n}\right\|_{\infty}\right) \\
& \geq \varepsilon
\end{aligned}
$$

This inequality, however, contradicts to the continuity of $f$ (note that $\lim _{n \rightarrow \infty} y_{n}=x$ ), and therefore, $B$ is equicontinuous at every $x \in K$.

Assume now that $B$ satisfies the properties from assertion (ii). In order to show that $B$ is compact, it suffices to show that every sequence $\left(f_{n}\right) \subseteq B$ admits a convergent subsequence, that is, $B$ is sequentially compact. So let $\left(f_{n}\right) \subseteq B$ be an arbitrary sequence.

Recall that every compact metric space is separable. Hence, there exists a sequence $\left(x_{m}\right)_{m \geq 1} \subseteq K$ which is dense in $K$.

Consider the sequence $\left(f_{n}\left(x_{1}\right)\right) \subseteq B_{x_{1}} \subseteq X$. Since $B_{x_{1}}$ is compact by assumption, there exists a subsequence $\left(f_{\varphi_{1}(n)}\right)$ of $\left(f_{n}\right)$ such that $\lim _{n \rightarrow \infty} f_{\varphi_{1}(n)}\left(x_{1}\right)$ exists.

Consider next the sequence $\left(f_{\varphi_{1}(n)}\left(x_{2}\right)\right) \subseteq B_{x_{2}} \subseteq X$. Since $B_{x_{2}}$ is compact by assumption, there exists a subsequence $\left(f_{\varphi_{2}(n)}\right)$ of $\left(f_{\varphi_{1}(n)}\right)$ such that $\lim _{n \rightarrow \infty} f_{\varphi_{2}(n)}\left(x_{2}\right)$ exists. Note that we have also the existence of the limit $\lim _{n \rightarrow \infty} f_{\varphi_{2}(n)}\left(x_{1}\right)$.

Iterating this argument, we obtain for every $m \geq 2$ a subsequence $\left(f_{\varphi_{m}(n)}\right)$ of $\left(f_{\varphi_{m-1}(n)}\right)$ such that $\lim _{n \rightarrow \infty} f_{\varphi_{m}(n)}\left(x_{i}\right)$ exists for every $1 \leq i \leq m$. These subsequences converge therefore pointwise at a finite number of elements of $K$.

We now consider the diagonal subsequence $\left(f_{\varphi(n)}\right)=\left(f_{\varphi_{n}(n)}\right)$. This diagonal subsequence has the property of being a subsequence of $\left(f_{\varphi_{m}(n)}\right)$ for every $m \geq 1$, up
to a finite number of initial elements perhaps. It enjoys therefore the property that $\lim _{n \rightarrow \infty} f_{\varphi(n)}\left(x_{m}\right)$ exists for every $m \geq 1$, that is, it converges pointwise on a dense subset of $K$. We will show that $\left(f_{\varphi(n)}\right)$ converges everywhere and uniformly on $K$. Since $C(K ; X)$ is complete, it suffices to show that $\left(f_{\varphi(n)}\right)$ is a Cauchy sequence in $C(K ; X)$.

Let $\varepsilon>0$. Since $B$ is equicontinuous at every $x \in K$, for every $x \in K$ there exists $\delta_{x}>0$ such that for every $y \in K$ and every $f \in B$ the implication

$$
\begin{equation*}
d(x, y)<\delta \quad \Rightarrow \quad\|f(x)-f(y)\|<\varepsilon \tag{B.2}
\end{equation*}
$$

is true. We clearly have $K=\bigcup_{x \in K} B\left(x, \delta_{x}\right)$, and since $K$ is compact, we find finitely many points $x_{1}^{\prime}, \ldots, x_{k}^{\prime}$ such that $K=\bigcup_{i=1}^{k} B\left(x_{i}^{\prime}, \delta_{i}\right)$ (with $\delta_{i}=\delta_{x_{i}^{\prime}}$. Since the sequence $\left(x_{m}\right)$ is dense in $K$, for every $1 \leq i \leq k$ there exists $m_{i} \geq 1$ such that $x_{m_{i}} \in B\left(x_{i}^{\prime}, \delta_{i}\right)$. Since the sequence $\left(f_{\varphi(n)}\right)$ converges pointwise on $\left(x_{m}\right)$, there exists $n_{0} \geq 0$ such that

$$
\text { for every } n, n^{\prime} \geq n_{0} \text { and every } 1 \leq i \leq k \quad\left\|f_{\varphi(n)}\left(x_{m_{i}}\right)-f_{\varphi\left(n^{\prime}\right)}\left(x_{m_{i}}\right)\right\|<\varepsilon
$$

Let now $x \in K$ be arbitrary. Then $x \in B\left(x_{i}, \delta_{i}\right)$ for some $1 \leq i \leq k$. Hence, for every $n, n^{\prime} \geq n_{0}$, by the preceding estimate and by the implication (B.2),

$$
\begin{aligned}
\left\|f_{\varphi(n)}(x)-f_{\varphi\left(n^{\prime}\right)}(x)\right\| \leq & \left\|f_{\varphi(n)}(x)-f_{\varphi(n)}\left(x_{i}^{\prime}\right)\right\|+ \\
& +\left\|f_{\varphi(n)}\left(x_{i}^{\prime}\right)-f_{\varphi(n)}\left(x_{m_{i}}\right)\right\|+ \\
& +\left\|f_{\varphi(n)}\left(x_{m_{i}}\right)-f_{\varphi\left(n^{\prime}\right)}\left(x_{m_{i}}\right)\right\|+ \\
& +\left\|f_{\varphi\left(n^{\prime}\right)}\left(x_{m_{i}}\right)-f_{\varphi\left(n^{\prime}\right)}\left(x_{i}^{\prime}\right)\right\|+ \\
& +\left\|f_{\varphi\left(n^{\prime}\right)}\left(x_{i}^{\prime}\right)-f_{\varphi\left(n^{\prime}\right)}(x)\right\| \\
\leq & 5 \varepsilon
\end{aligned}
$$

Since $n_{0} \geq 0$ did not depend on $x \in K$, and since $\varepsilon>0$ was arbitrary, this proves that $\left(f_{\varphi(n)}\right)$ is a Cauchy sequence in $C(K ; X)$. We have therefore proved that every sequence in $B$ admits a convergent subsequence. Since $B$ is closed, we obtain that $B$ is sequentially compact, and hence compact.

## Appendix C Calculus on Banach spaces

## C. 1 Differentiable functions between Banach spaces

Definition C.1. Let $X, Y$ be two Banach spaces, and let $U \subset X$ be open. A function $f: U \rightarrow Y$ is
a) differentiable at $x \in U$ if there exists a bounded linear operator $T \in \mathscr{L}(X, Y)$ such that

$$
\begin{equation*}
\lim _{\|h\| \rightarrow 0} \frac{f(x+h)-f(x)-T h}{\|h\|}=0 \tag{C.1}
\end{equation*}
$$

b) differentiable if it is differentiable at every point $x \in U$.

If $f$ is differentiable at a point $x \in U$, then $T \in \mathscr{L}(X, Y)$ is uniquely determined. We write $D f(x):=f^{\prime}(x):=T$ and call $D f(x)=f^{\prime}(x)$ the derivative of $f$ at $x$.

Lemma C.2. If a function $f: U \rightarrow Y$ is differentiable at $x \in U$, then it is continuous at $x$. In particular, every differentiable function is continuous.

Proof. Let $\left(x_{n}\right) \subset U$ be convergent to $x$. By definition (equation (C.1)) and continuity of $f^{\prime}(x)$,

$$
\begin{aligned}
\left\|f\left(x_{n}\right)-f(x)\right\| & \leq\left\|f\left(x_{n}\right)-f(x)-f^{\prime}(x)\left(x-x_{n}\right)\right\|+\left\|f^{\prime}(x)\left(x-x_{n}\right)\right\| \\
& \rightarrow 0
\end{aligned}
$$

as $n \rightarrow \infty$.
Definition C.3. Let $X, Y$ be two Banach spaces, and let $U \subset X$ be open. A function $f: U \rightarrow Y$ is called continuously differentiable if it is differentiable and if $f^{\prime}: U \rightarrow$ $\mathscr{L}(X, Y)$ is continuous. We denote by

$$
C^{1}(U ; Y):=\left\{f: U \rightarrow Y: f \text { differentiable and } f^{\prime} \in C(U ; \mathscr{L}(X, Y))\right\}
$$

the space of all continuously differentiable functions. Moreover, for $k \geq 2$, we denote by

$$
C^{k}(U ; Y):=\left\{f: U \rightarrow Y: f \text { differentiable and } f^{\prime} \in C^{k-1}(U ; \mathscr{L}(X, Y))\right\}
$$

the space of all $k$ times continuously differentiable functions.
Definition C.4. Let $X_{i}(1 \leq i \leq n)$ and $Y$ be Banach spaces. Let $U \subset \bigotimes_{i=1}^{n} X_{i}$ be open. We say that a function $f: U \rightarrow Y$ is at $a=\left(a_{i}\right)_{1 \leq i \leq n} \in U$ partially differentiably with respect to the $i$-th coordinate if the function

$$
f_{i}: U_{i} \subset X_{i} \rightarrow Y, \quad x_{i} \mapsto f\left(a_{1}, \ldots, x_{i}, \ldots, a_{n}\right)
$$

is differentiable in $a_{i}$. We write $\frac{\partial f}{\partial x_{i}}(a):=f_{i}^{\prime}\left(a_{i}\right) \in \mathscr{L}\left(X_{i}, Y\right)$.

## C. 2 Local inverse theorem and implicit function theorem

Let $X$ and $Y$ be two Banach spaces and let $U \subseteq X$ be an open subset. The following are two classical theorems in differential calculus.

Theorem C. 5 (Local inverse theorem). Let $f: U \rightarrow Y$ be continuously differentiable and $\bar{x} \in U$ such that $f^{\prime}(\bar{x}): X \rightarrow Y$ is an isomorphism, that is, bounded, bijective and the inverse is also bounded. Then there exist neighbourhoods $V \subset U$ of $\bar{x}$ and $W \subset Y$ of $f(\bar{x})$ such that $f: V \rightarrow W$ is a $C^{1}$ diffeomorphism, that is $f$ is continuously differentiable, bijective and the inverse $f^{-1}: W \rightarrow V$ is continuously differentiable, too.

Theorem C. 6 (Implicit function theorem). Assume that $X=X_{1} \times X_{2}$ for two Banach spaces $X_{1}, X_{2}$, and let $f: X \supset U \rightarrow Y$ be continuously differentiable. Let $\bar{x}=\left(\bar{x}_{1}, \bar{x}_{2}\right) \in U$ be such that $\frac{\partial f}{\partial x_{2}}(\bar{x}): X_{2} \rightarrow Y$ is an isomorphism. Then there exist neighbourhoods $U_{1} \subset X_{1}$ of $\bar{x}_{1}$ and $U_{2} \subset X_{2}$ of $\bar{x}_{2}, U_{1} \times U_{2} \subset U$, and a continuously differentiable function $g: U_{1} \rightarrow U_{2}$ such that

$$
\left\{x \in U_{1} \times U_{2}: f(x)=f(\bar{x})\right\}=\left\{\left(x_{1}, g\left(x_{1}\right)\right): x_{1} \in U_{1}\right\} .
$$

For the proof of the local inverse theorem, we need the following lemma.
Lemma C.7. Let $f: U \rightarrow Y$ be continuously differentiable such that $f: U \rightarrow f(U)$ is a homeomorphism, that is, continuous, bijective and with continuous inverse. Then $f$ is a $C^{1}$ diffeomorphism if and only if for every $x \in U$ the derivative $f^{\prime}(x): X \rightarrow Y$ is an isomorphism.

Proof. Assume first that $f$ is a $C^{1}$ diffeomorphism. When we differentiate the identities $x=f^{-1}(f(x))$ and $y=f\left(f^{-1}(y)\right)$, which are true for every $x \in U$ and every $y \in f(U)$, then we find

$$
\begin{aligned}
I_{X} & =\left(f^{-1}\right)^{\prime}(f(x)) f^{\prime}(x) \quad \text { for every } x \in U \text { and } \\
I_{Y} & =f^{\prime}\left(f^{-1}(y)\right)\left(f^{-1}\right)^{\prime}(y) \\
& =f^{\prime}(x)\left(f^{-1}\right)^{\prime}(f(x)) \quad \text { for every } x=f^{-1}(y) \in U
\end{aligned}
$$

As a consequence, $f^{\prime}(x)$ is an isomorphism for every $x \in U$.
For the converse, assume that $f^{\prime}(x)$ is an isomorphism for every $x \in U$. For every $x_{1}, x_{2} \in U$ one has, by differentiability,

$$
f\left(x_{2}\right)=f\left(x_{1}\right)+f^{\prime}\left(x_{1}\right)\left(x_{2}-x_{1}\right)+o\left(x_{2}-x_{1}\right),
$$

where $o$ depends on $x_{1}$ and $\lim _{x_{2} \rightarrow x_{1}} \frac{o\left(x_{2}-x_{1}\right)}{\left\|x_{2}-x_{1}\right\|}=0$. We have $x_{1}=f^{-1}\left(y_{1}\right)$ and $x_{2}=$ $f^{-1}\left(y_{2}\right)$ if we put $y_{i}:=f\left(x_{i}\right)$. Hence, the above identity becomes

$$
y_{2}=y_{1}+f^{\prime}\left(f^{-1}\left(y_{1}\right)\right)\left(f^{-1}\left(y_{2}\right)-f^{-1}\left(y_{1}\right)\right)+o\left(f^{-1}\left(y_{2}\right)-f^{-1}\left(y_{1}\right)\right) .
$$

To this identity, we apply the inverse operator $\left(f^{\prime}\left(f^{-1}\left(y_{1}\right)\right)\right)^{-1}$ and we obtain
$f^{-1}\left(y_{2}\right)=f^{-1}\left(y_{1}\right)+\left(f^{\prime}\left(f^{-1}\left(y_{1}\right)\right)\right)^{-1}\left(y_{2}-y_{1}\right)-\left(f^{\prime}\left(f^{-1}\left(y_{1}\right)\right)\right)^{-1} o\left(f^{-1}\left(y_{2}\right)-f^{-1}\left(y_{1}\right)\right)$.
Since $f^{-1}$ is continuous, the last term on the right-hand side of the last equality is sublinear. Hence, $f^{-1}$ is differentiable and

$$
\left(f^{-1}\right)^{\prime}\left(y_{1}\right)=\left(f^{\prime}\left(f^{-1}\left(y_{1}\right)\right)\right)^{-1} .
$$

From this identity (using that $f^{-1}$ and $f^{\prime}$ are continuous) we obtain that $f^{-1}$ is continuously differentiable. The claim is proved.

Proof (Proof of the local inverse theorem). Consider the function

$$
\begin{aligned}
g: U & \rightarrow X, \\
& x \mapsto f^{\prime}(\bar{x})^{-1} f(x) .
\end{aligned}
$$

It suffices to show that $g: V \rightarrow W$ is a $C^{1}$ diffeomorphism for appropriate neighbourhoods $V$ of $\bar{x}$ and $W$ of $g(\bar{x})$.

Consider also the function

$$
\begin{aligned}
\varphi: U & \rightarrow X, \\
x & \mapsto x-g(x) .
\end{aligned}
$$

This function $\varphi$ is continuously differentiable and $\varphi^{\prime}(x)=I-f^{\prime}(\bar{x})^{-1} f^{\prime}(x)$ for every $x \in U$. In particular, $\varphi^{\prime}(\bar{x})=0$. By continuity of $\varphi^{\prime}$, there exists $r>0$ and $L<1$ such that $\left\|\varphi^{\prime}(x)\right\| \leq L$ for every $x \in \bar{B}(\bar{x}, r) \subset U$. Hence,

$$
\left\|\varphi\left(x_{1}\right)-\varphi\left(x_{2}\right)\right\| \leq L\left\|x_{1}-x_{2}\right\| \quad \text { for every } x_{1}, x_{2} \in \bar{B}(\bar{x}, r) .
$$

By the definition of $\varphi$, this implies

$$
\begin{align*}
\left\|g\left(x_{1}\right)-g\left(x_{2}\right)\right\| & =\left\|x_{1}-x_{2}-\left(\varphi\left(x_{1}\right)-\varphi\left(x_{2}\right)\right)\right\|  \tag{C.2}\\
& \geq\left\|x_{1}-x_{2}\right\|-L\left\|x_{1}-x_{2}\right\| \\
& =(1-L)\left\|x_{1}-x_{2}\right\| .
\end{align*}
$$

We claim that for every $y \in \bar{B}(g(\bar{x}),(1-L) r)$ there exists a unique $x \in \bar{B}(\bar{x}, r)$ such that $g(x)=y$.

The uniqueness follows from (C.2).
In order to prove existence, let $x_{0}=\bar{x}$, and then define recursively $x_{n+1}=y+$ $\varphi\left(x_{n}\right)=y+x_{n}-f^{\prime}(\bar{x})^{-1} f\left(x_{n}\right)$ for every $n \geq 0$. Then

$$
\begin{aligned}
\left\|x_{n}-\bar{x}\right\| & =\left\|\sum_{k=0}^{n-1} x_{k+1}-x_{k}\right\| \\
& \leq\left\|x_{1}-x_{0}\right\|+\sum_{k=1}^{n-1}\left\|\varphi\left(x_{k}\right)-\varphi\left(x_{k-1}\right)\right\| \\
& \leq \sum_{k=0}^{n-1} L^{k}\left\|x_{1}-x_{0}\right\| \\
& =\frac{1-L^{n}}{1-L}\|y-g(\bar{x})\| \\
& \leq\left(1-L^{n}\right) r \leq r
\end{aligned}
$$

which implies $x_{n} \in \bar{B}(\bar{x}, r)$ for every $n \geq 0$. Similarly, for every $n \geq m \geq 0$,

$$
\left\|x_{n}-x_{m}\right\| \leq \sum_{k=m}^{n-1} L^{k}\|y-g(\bar{x})\|
$$

so that the sequence $\left(x_{n}\right)$ is a Cauchy sequence in $\bar{B}(\bar{x}, r)$. Since $\bar{B}(\bar{x}, r)$ is complete, there exists $\lim _{n \rightarrow \infty} x_{n}=: x \in \bar{B}(\bar{x}, r)$. By continuity,

$$
x=y+\varphi(x)=y+x-g(x)
$$

or

$$
g(x)=y .
$$

This proves the above claim, that is, $g$ is locally invertible. It remains to show that $g^{-1}$ is continuous (then $g$ is a homeomorphism, and therefore a $C^{1}$ diffeomorphism by Lemma C.7). Contiunity of the inverse function, however, is a direct consequence of (C.2) (which even implies Lipschitz continuity).

Remark C.8. The iteration formula

$$
x_{n+1}=y+x_{n}-f^{\prime}(\bar{x})^{-1} f\left(x_{n}\right)
$$

used in the proof of the local inverse theorem in order to find a solution of $g(x)=$ $f^{\prime}(\bar{x})^{-1} f(x)=y$ should be compared to the discrete Newton iteration

$$
x_{n+1}=y+x_{n}-f^{\prime}\left(x_{n}\right)^{-1} f\left(x_{n}\right) ;
$$

see Theorem C. 11 below.
Proof (Proof of the implicit function theorem). Consider the function

$$
\begin{aligned}
F: U & \rightarrow X_{1} \times Y \\
\left(x_{1}, x_{2}\right) & \mapsto\left(x_{1}, f\left(x_{1}, x_{2}\right)\right)
\end{aligned}
$$

Then $F$ is continuously differentiable and

$$
F^{\prime}(\bar{x})\left(h_{1}, h_{2}\right)=\left(h_{1}, \frac{\partial f}{\partial x_{1}}(\bar{x}) h_{1}+\frac{\partial f}{\partial x_{2}}(\bar{x}) h_{2}\right) .
$$

In particular, by the assumption, $F^{\prime}(\bar{x})$ is locally invertible with inverse

$$
F^{\prime}(\bar{x})^{-1}\left(y_{1}, y_{2}\right)=\left(y_{1},\left(\frac{\partial f}{\partial x_{2}}(\bar{x})\right)^{-1}\left(y_{2}-\frac{\partial f}{\partial x_{1}}(\bar{x}) y_{1}\right)\right)
$$

By the local inverse theorem (Theorem C.5), there exists a neighbourhood $U_{1}$ of $\bar{x}_{1}$, a neighbourhood $U_{2}$ of $\bar{x}_{2}$ and a neighbourhood $V$ of $\left(\bar{x}_{1}, f(\bar{x})\right)=F(\bar{x})$ such that $F: U_{1} \times U_{2} \rightarrow V$ is a $C^{1}$ diffeomorphism. The inverse is of the form

$$
F^{-1}\left(y_{1}, y_{2}\right)=\left(y_{1}, h_{2}\left(y_{1}, y_{2}\right)\right)
$$

where $h_{2}$ is a function such that $f\left(y_{1}, h_{2}\left(y_{1}, y_{2}\right)\right)=y_{2}$. Let

$$
\tilde{U}_{1}:=\left\{x_{1} \in U_{1}:\left(x_{1}, f(\bar{x})\right) \in V\right\}
$$

Then $\tilde{U}_{1}$ is open by continuity of the function $x_{1} \mapsto\left(x_{1}, f(\bar{x})\right)$, and $\bar{x}_{1} \in \tilde{U}_{1}$. We restrict $F$ to $\tilde{U}_{1} \times U_{2}$, and we define

$$
\begin{align*}
g: \tilde{U}_{1} & \rightarrow X_{2}  \tag{C.3}\\
x_{1} & \mapsto g\left(x_{1}\right)=F^{-1}\left(x_{1}, f(\bar{x})\right)_{2}
\end{align*}
$$

where $F^{-1}(\cdot)_{2}$ denotes the second component of $F^{-1}(\cdot)$. Then $g$ is continuously differentiable, $g\left(\tilde{U}_{1}\right) \subset U_{2}$ and $g$ satisfies the required property of the implicit function.

Lemma C. 9 (Higher regularity of the local inverse). Let $f \in C^{k}(U ; Y)$ for some $k \geq 1$ and assum that $f: U \rightarrow f(U)$ is a $C^{1}$ diffeomorphism. Then $f$ is a $C^{k}$ diffeomorphism, that is, $f^{-1}$ is $k$ times continuously differentiable.

Proof. For every $y \in f(U)$ we have

$$
\left(f^{-1}\right)^{\prime}(y)=f^{\prime}\left(f^{-1}(y)\right)^{-1}
$$

The proof therefore follows by induction on $k$.
Lemma C. 10 (Higher regularity of the implicit function). If, in the implicit function theorem (Theorem C.6), the function $f$ is $k$ times continuously differentiable, then the implicit function $g$ is also $k$ times continuously differentiable.

Proof. This follows from the previous lemma (Lemma C.9) and the definition of the implicit function in the proof of the implicit function theorem.

## C. 3 * Newton's method

Theorem C. 11 (Newton's method). Let $X$ and $Y$ be two Banach spaces, $U \subset X$ an open set. Let $f \in C^{1}(U ; Y)$ and assume that there exists $\bar{x} \in U$ such that (i) $f(\bar{x})=0$ and (ii) $f^{\prime}(\bar{x}) \in \mathscr{L}(X, Y)$ is an isomorphism. Then there exists a neighbourhood $V \subset U$ of $\bar{x}$ such that for every $x_{0} \in V$ the operator $f^{\prime}\left(x_{0}\right)$ is an isomorphism, the sequence $\left(x_{n}\right)$ defined iteratively by

$$
\begin{equation*}
x_{n+1}=x_{n}-f^{\prime}\left(x_{n}\right)^{-1} f\left(x_{n}\right), \quad n \geq 0 \tag{C.4}
\end{equation*}
$$

remains in $V$ and $\lim _{n \rightarrow \infty} x_{n}=\bar{x}$.
Proof. By Corollary B. 42 and continuity, there exists a neighbourhood $\tilde{V} \subset U$ of $\bar{x}$ such that $f^{\prime}(x)$ is isomorphic for all $x \in \tilde{V}$. Next, it will be useful to define the auxiliary function $\varphi: \tilde{V} \rightarrow X$ by

$$
\varphi(x):=x-f^{\prime}(x)^{-1} f(x), \quad x \in \tilde{V} .
$$

Since $f(\bar{x})=0$, we find that for every $x \in \tilde{V}$

$$
\begin{aligned}
\varphi(x)-\varphi(\bar{x}) & =x-f^{\prime}(x)^{-1}(f(x)-f(\bar{x}))-\bar{x} \\
& =x-\bar{x}-f^{\prime}(x)^{-1}\left(f^{\prime}(\bar{x})(x-\bar{x})+r(x-\bar{x})\right),
\end{aligned}
$$

so that by the continuity of $f^{\prime}(\cdot)^{-1}$

$$
\lim _{x \rightarrow \bar{x}} \frac{\|\varphi(x)-\varphi(\bar{x})\|}{\|x-\bar{x}\|}=0 .
$$

Hence, there exists $r>0$ such that $V:=B(\bar{x}, r) \subset \tilde{V} \subset U$ and such that for every $x \in V$

$$
\|\varphi(x)-\bar{x}\|=\|\varphi(x)-\varphi(\bar{x})\| \leq \frac{1}{2}\|x-\bar{x}\| .
$$

This implies that for every $x_{0} \in V$ one has $\varphi\left(x_{0}\right) \in V$ and if we define iteratively $x_{n+1}=\varphi\left(x_{n}\right)=\varphi^{n+1}\left(x_{0}\right)$, then

$$
\left\|x_{n}-\bar{x}\right\| \leq\left(\frac{1}{2}\right)^{n}\left\|x_{0}-\bar{x}\right\| \rightarrow 0 \text { as } n \rightarrow \infty
$$

## Appendix D <br> Hilbert spaces

Let $H$ be a vector space over $\mathbb{K}$.

## D. 1 Inner product spaces

Definition D.1. A function $\langle\cdot, \cdot\rangle: H \times H \rightarrow \mathbb{K}$ is called an inner product if for every $x, y, z \in H$ and every $\lambda \in \mathbb{K}$
(i) $\langle x, x\rangle \geq 0$ for every $x \in H$ and $\langle x, x\rangle=0$ if and only if $x=0$,
(ii) $\langle x, y\rangle=\overline{\langle y, x\rangle}$,
(iii) $\langle\lambda x+y, z\rangle=\lambda\langle x, z\rangle+\langle y, z\rangle$.

A pair $(H,\langle\cdot, \cdot\rangle)$ of a vector space over $\mathbb{K}$ and a scalar product is called an inner product space.

Example D.2. 1. On the space $H=\mathbb{K}^{d}$,

$$
\langle x, y\rangle:=\sum_{i=1}^{d} x_{i} \bar{y}_{i}
$$

defines an inner product.
2. On the space $H=l^{2}:=\left\{\left(x_{n}\right) \subset \mathbb{K}: \sum\left|x_{n}\right|^{2}<\infty\right\}$,

$$
\langle x, y\rangle:=\sum_{n} x_{n} \overline{y_{n}}
$$

defines an inner product.
3. On the space $H=C([0,1])$, the Riemann integral

$$
\langle f, g\rangle:=\int_{0}^{1} f(x) \overline{g(x)} d x
$$

defines an inner product.
4. On the space $H=L^{2}(\Omega)$, the integral

$$
\langle f, g\rangle:=\int_{\Omega} f \bar{g} d \mu
$$

defines an inner product.
Lemma D.3. Let $\langle\cdot, \cdot\rangle$ be an inner product on a vector space $H$. Then, for every $x$, $y, z \in H$ and $\lambda \in \mathbb{K}$
(iv) $\langle x, \lambda y+z\rangle=\bar{\lambda}\langle x, y\rangle+\langle x, z\rangle$.

Proof.

$$
\langle x, \lambda y+z\rangle=\overline{\langle\lambda y+z, x\rangle}=\bar{\lambda} \overline{\langle y, x\rangle}+\overline{\langle z, x\rangle}=\bar{\lambda}\langle x, y\rangle+\langle x, z\rangle
$$

In the following, if $H$ is an inner product space, then we put

$$
\|x\|:=\sqrt{\langle x, x\rangle}, \quad x \in H
$$

Lemma D. 4 (Cauchy-Schwarz inequality). Let $H$ be an inner product space. Then, for every $x, y \in H$,

$$
|\langle x, y\rangle| \leq\|x\|\|y\|
$$

and equality holds if and only if $x$ and $y$ are colinear.
Proof. Let $\lambda \in \mathbb{K}$. Then

$$
\begin{aligned}
0 & \leq\langle x+\lambda y, x+\lambda y\rangle \\
& =\langle x, x\rangle+\langle\lambda y, x\rangle+\langle x, \lambda y\rangle+|\lambda|^{2}\langle y, y\rangle \\
& =\langle x, x\rangle+\lambda \overline{\langle x, y\rangle}+\bar{\lambda}\langle x, y\rangle+|\lambda|^{2}\langle y, y\rangle
\end{aligned}
$$

that is,

$$
\begin{equation*}
0 \leq\|x+\lambda y\|^{2}=\|x\|^{2}+2 \operatorname{Re} \bar{\lambda}\langle x, y\rangle+|\lambda|^{2}\|y\|^{2} \tag{D.1}
\end{equation*}
$$

Assuming that $y \neq 0$ (for $y=0$ the Cauchy-Schwarz inequality is trivial), we may put $\lambda:=-\langle x, y\rangle /\|y\|^{2}$. Then

$$
\begin{aligned}
0 & \leq\left\langle x-\frac{\langle x, y\rangle}{\|y\|^{2}} y, x-\frac{\langle x, y\rangle}{\|y\|^{2}} y\right\rangle \\
& =\|x\|^{2}-\frac{|\langle x, y\rangle|^{2}}{\|y\|^{2}}
\end{aligned}
$$

which is the Cauchy-Schwarz inequality. The calculation also shows that equality holds if and only if $x=\lambda y$, that is, if $x$ and $y$ are colinear.

Lemma D.5. Every inner product space $H$ is a normed linear space for the norm

$$
\|x\|=\sqrt{\langle x, x\rangle}, \quad x \in H
$$

Proof. Properties (i) and (ii) in the definition of a norm follow from the properties (i) and (iii) (together with Lemma D.3) in the definition of an inner product. The only difficulty is to show that $\|\cdot\|$ satisfies the triangle inequality. This, however, follows from putting $\lambda=1$ in (D.1) and estimating with the Cauchy-Schwarz inequality:

$$
\|x+y\|^{2} \leq(\|x\|+\|y\|)^{2} .
$$

Definition D.6. A complete inner product space is called a Hilbert space.
Example D.7. The spaces $\mathbb{K}^{d}$ (with euclidean inner product), $l^{2}$ and $L^{2}(\Omega)$ are Hilbert spaces. More examples are given by the Sobolev spaces defined below.

Lemma D. 8 (Completion of an inner product space). Let $H$ be an inner product space. Then there exists a Hilbert space $K$ and a bounded linear operator $j: H \rightarrow K$ such that for every $x, y \in H$

$$
\langle x, y\rangle_{H}=\langle j(x), j(y)\rangle_{K},
$$

and such that $j(H)$ is dense in $K$. The Hilbert space $K$ is unique up to isometry. It is called the completion of $H$.

Lemma D. 9 (Parallelogram identity). Let $H$ be an inner product space. Then for every $x, y \in H$

$$
\|x+y\|^{2}+\|x-y\|^{2}=2\left(\|x\|^{2}+\|y\|^{2}\right)
$$

Proof. The parallelogram identity follows immediately from (D.1) by putting $\lambda=$ $\pm 1$ and adding up.

Exercise D. 10 (von Neumann) Show that a norm satisfying the parallelogram identity comes from a scalar product. That means, the parallelogram identity characterises inner product spaces.

Definition D.11. A subset $K$ of a vector space $X$ (over $\mathbb{K}$ ) is convex if for every $x$, $y \in K$ and every $t \in[0,1]$ one has $t x+(1-t) y \in K$.

Theorem D. 12 (Projection onto closed, convex sets). Given a nonempty closed, convex subset $K$ of a Hilbert space $H$, and given a point $x \in H$, there exists a unique $y \in K$ such that

$$
\|x-y\|=\inf \{\|x-z\|: z \in K\} .
$$

Proof. Let $d:=\inf \{\|x-z\|: z \in K\}$, and choose $\left(y_{n}\right) \in K$ such that

$$
\begin{equation*}
\lim _{n \rightarrow \infty}\left\|x-y_{n}\right\|=d \tag{D.2}
\end{equation*}
$$

Applying the parallelogram identity to $\left(x-y_{n}\right) / 2$ and $\left(x-y_{m}\right) / 2$, we obtain

$$
\left\|x-\frac{y_{n}+y_{m}}{2}\right\|^{2}+\frac{1}{4}\left\|y_{n}-y_{m}\right\|^{2}=\frac{1}{2}\left(\left\|x-y_{n}\right\|^{2}+\left\|x-y_{m}\right\|^{2}\right) .
$$

Since $K$ is convex, $\frac{y_{n}+y_{m}}{2} \in K$ and hence $\left\|x-\frac{y_{n}+y_{m}}{2}\right\|^{2} \geq d^{2}$. Using this and (D.2), the last identity implies that $\left(y_{n}\right)$ is a Cauchy sequence. Since $H$ is complete, $y:=$ $\lim _{n \rightarrow \infty} y_{n}$ exists. Since $K$ is closed, $y \in K$. Moreover, $\|x-y\|=\lim _{n \rightarrow \infty}\left\|x-y_{n}\right\|=d$, so that $y$ is a minimizer for the distance to $x$. To see that there is only on such minimizer, suppose that $y^{\prime} \in K$ is a second one, and apply the parallelogram identity to $x-y$ and $x-y^{\prime}$.

Definition D.13. Let $H$ be an inner product space. We say that two vectors $x, y \in H$ are orthogonal (and we write $x \perp y$ ), if $\langle x, y\rangle=0$. Given a subset $S \subset H$, we define the orthogonal space $S^{\perp}:=\{y \in H: x \perp y$ for all $x \in S\}$. If $S=K$ is a linear subspace of $H$, then we call $K^{\perp}$ also the orthogonal complement of $K$.

Theorem D.14. Let $H$ be a Hilbert space, $S \subset H$ be a subset and $K$ a closed linear subspace. Then:
a) $S^{\perp}$ is a closed linear subspace of $H$,
b) $K$ and $K^{\perp}$ are complementary subspaces, i.e. every $x \in H$ can be decomposed uniquely as a sum of an $x_{0} \in K$ and an $x_{1} \in K^{\perp}$,
c) $\left(K^{\perp}\right)^{\perp}=K$ and $\left(S^{\perp}\right)^{\perp}=\overline{\operatorname{span}} S$.
d) $\operatorname{span} S$ is dense in $H$ if and only if $S^{\perp}=\{0\}$.

Proof. (a) It follows from the bilinearity of the inner product that $S^{\perp}$ is a linear subspace of $H$. Let $\left(y_{n}\right) \in S^{\perp}$ be convergent to some $y \in H$. Then, for every $x \in S$, by the Cauchy-Schwarz inequality,

$$
\langle x, y\rangle=\lim _{n \rightarrow \infty}\left\langle x, y_{n}\right\rangle=0
$$

that is, $y \in S^{\perp}$ and therefore $S^{\perp}$ is closed.
(b) For every $x \in H$ we let $x_{0} \in K$ be the unique element (Theorem D.12) such that

$$
\left\|x-x_{0}\right\|=\inf \{\|x-y\|: y \in K\}
$$

Put $x_{1}=x-x_{0}$. For every $y \in K$ and every $\lambda \in \mathbb{K}$, by the minimum property of $x_{0}$,

$$
\begin{aligned}
\left\|x_{1}\right\|^{2} & \leq\left\|x_{1}-\lambda y\right\|^{2} \\
& =\left\|x_{1}\right\|^{2}-2 \operatorname{Re} \bar{\lambda}\left\langle x_{1}, y\right\rangle+|\lambda|^{2}\|y\|^{2} .
\end{aligned}
$$

This implies that $\left\langle x_{1}, y\right\rangle=0$, that is, $x_{1} \in K^{\perp}$. Every decomposition $x=x_{0}+x_{1}$ with $x_{0} \in K$ and $x_{1} \in K^{\perp}$ is unique since $x \in K \cap K^{\perp}$ implies $\langle x, x\rangle=0$, that is, $x=0$.
(c) and (d) follow immediately from (a) and (b).

Lemma D. 15 (Pythagoras). Let $H$ be an inner product space. Whenever $x, y \in H$ are orthogonal, then

$$
\|x+y\|^{2}=\|x\|^{2}+\|y\|^{2} .
$$

Proof. The claim follows from (D.1) and putting $\lambda=1$.

Definition D.16. Let $X$ be a normed space. We call an operator $P: X \rightarrow X$ a projection if $P^{2}=P$.

Lemma D.17. Let $X$ be a normed space and let $P \in \mathscr{L}(X)$ be a bounded projection. Then the following are true:
a) $Q=I-P$ is a projection.
b) Either $P=0$ or $\|P\| \geq 1$.
c) The kernel $\operatorname{ker} P$ and the range $\operatorname{ran} P$ are closed in $X$.
d) Every $x \in X$ can be decomposed uniquely as a sum of an $x_{0} \in \operatorname{ker} P$ and an $x_{1} \in \operatorname{ran} P$, and $X \cong \operatorname{ker} P \oplus \operatorname{ran} P$.

Proof. (a) $Q^{2}=(I-P)^{2}=I-2 P+P^{2}=I-P=Q$.
(b) follows from $\|P\|=\left\|P^{2}\right\| \leq\|P\|^{2}$.
(c) Since $\{0\}$ is closed in $X$ and since $P$ is continuous, $\operatorname{ker} P=P^{-1}(\{0\})$ is closed. Similarly, $\operatorname{ran} P=\operatorname{ker}(I-P)$ is closed.
(d) For every $x \in X$ we can write $x=P x+(I-P) x=x_{1}+x_{2}$ with $x_{1} \in \operatorname{ran} P$ and $x_{2} \in \operatorname{ker} P$. The decomposition is unique since if $x \in \operatorname{ker} P \cap \operatorname{ran} P$, then $x=P x=0$. This proves that the vector spaces $X$ and $\operatorname{ker} P \oplus \operatorname{ran} P$ are isomorphic. That they are also isomorphic as normed spaces follows from the continuity of $P$.

Lemma D.18. Let $H$ be a Hilbert space and $K \subset H$ be a closed linear subspace. For every $x \in H$ we let $x_{1}=P x$ be the unique element in $K$ which minimizes the distance to $x$ (Theorem D.12). Then $P: H \rightarrow H$ is a bounded projection satisfying $\operatorname{ran} P=K$. Moreover, $\operatorname{ker} P=K^{\perp}$. We call $P$ the orthogonal projection onto $K$.

## D. 2 Orthogonal decomposition

Definition D.19. We call a metric space separable if there exists a countable dense subset.

Example D.20. The space $\mathbb{R}^{d}$ (or $\mathbb{C}^{d}$ ) is separable: one may take $\mathbb{Q}^{d}$ as an example of a dense countable subset. It is not too difficult to see that subsets of separable metric spaces are separable (note, however, that in general the dense subset has to be constructed carefully), and that finite products of separable metric spaces are separable.

Lemma D.21. A normed space $X$ is separable if and only if there exists a sequence $\left(x_{n}\right) \subset X$ such that $\operatorname{span}\left\{x_{n}: n \in \mathbb{N}\right\}$ is dense in $X$ (such a sequence is in general called $a$ total sequence).

Proof. If $X$ is separable, then there exists a sequence $\left(x_{n}\right) \subset X$ such that $\left\{x_{n}: n \in \mathbb{N}\right\}$ is dense. In particular, the larger set span $\left\{x_{n}: n \in \mathbb{N}\right\}$ is dense.

If, one the other hand, there exists a total sequence $\left(x_{n}\right) \subset X$, and if we put $D=\mathbb{Q}$ in the case $\mathbb{K}=\mathbb{R}$ and $D=\mathbb{Q}+i \mathbb{Q}$ in the case $\mathbb{K}=\mathbb{C}$, then the set

$$
\left\{\sum_{i=1}^{m} \lambda_{i} x_{n_{i}}: m \in \mathbb{N}, \lambda_{i} \in D, n_{i} \in \mathbb{N}\right\}
$$

is dense in $X$ (in fact, the closure contains all finite linear combinations of the $x_{n}$, that is, it contains $\left.\operatorname{span}\left\{x_{n}: n \in \mathbb{N}\right\}\right)$. It is an exercise to show that this set is countable. The claim follows.

Corollary D.22. The space $\left(C([0,1]),\|\cdot\|_{\infty}\right)$ is separable.
Proof. By Weierstrass' theorem, the subspace of all polynomials is dense in $C([0,1])$ (Weierstrass' theorem says that every continuous function $f:[0,1] \rightarrow \mathbb{R}$ can be uniformly approximated by polynomials). The polynomials, however, are the linear span of the monomials $f_{n}(t)=t^{n}$. The claim therefore follows from Lemma D. 21 .

Corollary D.23. The space $l^{p}$ is separable if $1 \leq p<\infty$. The space $c_{0}$ is separable.
Proof. Let $e_{n}=\left(\delta_{n k}\right)_{k} \in l^{p}$ be the $n$-th unit vector in $l^{p}$ (here $\delta_{n k}$ denotes the Kronecker symbol: $\delta_{n k}=1$ if $n=k$ and $\delta_{n k}=0$ otherwise). Then $\operatorname{span}\left\{e_{n}: n \in \mathbb{N}\right\}=c_{00}$ (the space of all finite sequences) is dense in $l^{p}$ if $1 \leq p<\infty$. The claim for $l^{p}$ follows from Lemma D.21. The argument for $c_{0}$ is similar.

Lemma D.24. The space $l^{\infty}$ is not separable.
Proof. The set $\{0,1\}^{\mathbb{N}} \subset l^{\infty}$ of all sequences taking only values 0 or 1 is uncountable. Moreover, whenever $x, y \in\{0,1\}^{\mathbb{N}}, x \neq y$, then

$$
\|x-y\|_{\infty}=1
$$

Hence, the balls $B\left(x, \frac{1}{2}\right)$ with centers $x \in\{0,1\}^{\mathbb{N}}$ and radius $\frac{1}{2}$ are mutually disjoint. If $l^{\infty}$ was separable, that is, if there exists a dense countable set $D \subset l^{\infty}$, then in each $B\left(x, \frac{1}{2}\right)$ there exists at least one element $y \in D$, a contradiction.

Definition D.25. Let $H$ be an inner product space. A family $\left(e_{l}\right)_{l \in I} \subset H$ is called
a) an orthogonal system if $\left(e_{l}, e_{k}\right)=0$ whenever $l \neq k$,
b) an orthonormal system if it is an orthogonal system and $\left\|e_{l}\right\|=1$ for every $l \in I$, and
c) an orthonormal basis if it is an orthonormal system and $\operatorname{span}\left\{e_{l}: l \in I\right\}$ is dense in $H$.

Lemma D. 26 (Gram-Schmidt process). Let $\left(x_{n}\right)$ be a sequence in an inner product space $H$. Then there exists an orthonormal system $\left(e_{n}\right)$ such that $\operatorname{span}\left\{x_{n}\right\}=$ $\operatorname{span}\left\{e_{n}\right\}$.

Proof. Passing to a subsequence, if necessary, we may assume that the $\left(x_{n}\right)$ are linearly independent.

Let $e_{1}:=x_{1} /\left\|x_{1}\right\|$. Then $e_{1}$ and $x_{1}$ span the same linear subspace. Next, assume that we have constructed an orthonormal system $\left(e_{k}\right)_{1 \leq k \leq n}$ such that

$$
\operatorname{span}\left\{x_{k}: 1 \leq k \leq n\right\}=\operatorname{span}\left\{e_{k}: 1 \leq k \leq n\right\}
$$

Let $e_{n+1}^{\prime}:=x_{n+1}-\sum_{k=1}^{n}\left\langle x_{n+1}, e_{k}\right\rangle e_{k}$. Since the $x_{n}$ are linearly independent, we find $e_{n+1}^{\prime} \neq 0$. Let $e_{n+1}:=e_{n+1}^{\prime} /\left\|e_{n+1}^{\prime}\right\|$. By construction, for every $1 \leq k \leq n$, $\left\langle e_{n+1}, e_{k}\right\rangle=0$, and

$$
\operatorname{span}\left\{x_{k}: 1 \leq k \leq n+1\right\}=\operatorname{span}\left\{e_{k}: 1 \leq k \leq n+1\right\}
$$

Proceeding inductively, the claim follows.
Corollary D.27. Every separable inner product space admits an orthonormal basis.
Example D.28. Consider the inner product space $C([-1,1])$ equiped with the scalar product $\langle f, g\rangle=\int_{-1}^{1} f(t) \overline{g(t)} d t$ and resulting norm $\|\cdot\|_{2}$. Let $f_{n}(t):=t^{n}(n \geq 0)$, so that span $\left\{f_{n}\right\}$ is the space of all polynomials on the interval $[-1,1]$. Applying the Gram-Schmidt process to the sequence $\left(f_{n}\right)$ yields a orthonormal sequence $\left(p_{n}\right)$ of polynomials. The $p_{n}$ are called Legendre polynomials.

Recall that the space of all polynomials is dense in $C([-1,1])$ by Weierstrass' theorem (even for the uniform norm; a fortiori also for the norm $\|\cdot\|_{2}$ ). Hence, the Legendre polynomials form an orthonormal basis in $C([-1,1])$.

Lemma D. 29 (Bessel's inequality). Let $H$ be an inner product space, $\left(e_{n}\right)_{n \in \mathbb{N}} \subset H$ an orthonormal system. Then, for every $x \in H$,

$$
\sum_{n \in \mathbb{N}}\left|\left\langle x, e_{n}\right\rangle\right|^{2} \leq\|x\|^{2}
$$

Proof. Let $N \in \mathbb{N}$. Put $x_{N}=x-\sum_{n=1}^{N}\left\langle x, e_{n}\right\rangle e_{n}$ so that $x_{N} \perp e_{n}$ for every $1 \leq n \leq N$. By Pythagoras (Lemma D.15),

$$
\begin{aligned}
\|x\|^{2} & =\left\|x_{N}\right\|^{2}+\left\|\sum_{n=1}^{N}\left\langle x, e_{n}\right\rangle e_{n}\right\|^{2} \\
& =\left\|x_{N}\right\|^{2}+\sum_{n=1}^{N}\left|\left\langle x, e_{n}\right\rangle\right|^{2} \\
& \geq \sum_{n=1}^{N}\left|\left\langle x, e_{n}\right\rangle\right|^{2}
\end{aligned}
$$

Since $N$ was arbitrary, the claim follows.
Lemma D.30. Let $H$ be a (separable) Hilbert space, $\left(e_{n}\right)_{n \in \mathbb{N}} \subset H$ an orthonormal system. Then:
a) For every $x \in H$, the series $\sum_{n \in \mathbb{N}}\left\langle x, e_{n}\right\rangle e_{n}$ converges.
b) $P: H \rightarrow H, x \mapsto \sum_{n \in \mathbb{N}}\left\langle x, e_{n}\right\rangle e_{n}$ is the orthogonal projection onto $\overline{\operatorname{span}}\left\{e_{n}: n \in\right.$ $\mathbb{N}\}$.

Proof. (a) Let $x \in H$. Since $\left(e_{n}\right)$ is an orthonormal system, by Pythagoras (Lemma D.15), for every $l>k \geq 1$,

$$
\begin{aligned}
\left\|\sum_{n=1}^{l}\left\langle x, e_{n}\right\rangle e_{n}-\sum_{n=1}^{k}\left\langle x, e_{n}\right\rangle e_{n}\right\|^{2} & =\left\|\sum_{n=k+1}^{l}\left\langle x, e_{n}\right\rangle e_{n}\right\|^{2} \\
& =\sum_{n=k+1}^{l}\left|\left\langle x, e_{n}\right\rangle\right|^{2}
\end{aligned}
$$

Hence, by Bessel's inequality, the sequence $\left(\sum_{n=1}^{l}\left\langle x, e_{n}\right\rangle e_{n}\right)$ of partial sums forms a Cauchy sequence. Since $H$ is complete, the series $\sum_{n \in \mathbb{N}}\left\langle x, e_{n}\right\rangle e_{n}$ converges.
(b) is an exercise.

Theorem D.31. Let H be a (separable) Hilbert space, $\left(e_{n}\right)_{n \in \mathbb{N}}$ an orthonormal system. Then the following are equivalent:
(i) $\left(e_{n}\right)_{n \in \mathbb{N}}$ is an orthonormal basis.
(ii) If $x \perp e_{n}$ for every $n \in \mathbb{N}$, then $x=0$.
(iii) $x=\sum_{n \in \mathbb{N}}\left\langle x, e_{n}\right\rangle e_{n}$ for every $x \in H$.
(iv) $\langle x, y\rangle=\sum_{n \in \mathbb{N}}\left\langle x, e_{n}\right\rangle\left\langle e_{n}, y\right\rangle$ for every $x, y \in H$.
(v) (Parseval's identity) For every $x \in H$,

$$
\|x\|^{2}=\sum_{n \in \mathbb{N}}\left|\left\langle x, e_{n}\right\rangle\right|^{2}
$$

Proof. (i) $\Rightarrow$ (ii) follows from Theorem D. 14 .
(ii) $\Rightarrow$ (iii) follows from Lemma D. 30 (i). In fact, let $x_{0}=\sum_{n \in \mathbb{N}}\left\langle x, e_{n}\right\rangle e_{n}$ (which exists by Lemma D. 30 (i)). Then $\left\langle x-x_{0}, e_{n}\right\rangle=0$ for every $n \in \mathbb{N}$, and by assumption (ii), this implies $x=x_{0}$.
(iii) $\Rightarrow$ (iv) follows when multiplying $x$ scalarly with $y$, applying also the CauchySchwarz inequality for the sequences $\left(\left\langle x, e_{l}\right\rangle\right),\left(\left\langle e_{l}, y\right\rangle\right) \in l^{2}$.
(iv) $\Rightarrow$ (v) follows from putting $x=y$.
(v) $\Rightarrow$ (i). Let $x \in \operatorname{span}\left\{e_{n}: n \in \mathbb{N}\right\}^{\perp}$. Then Parseval's identity implies $\|x\|^{2}=0$, that is, $x=0$. By Theorem D.14, $\operatorname{span}\left\{e_{n}: n \in \mathbb{N}\right\}$ is dense in $H$, that is, $\left(e_{n}\right)$ is an orthonormal basis.

Definition D.32. A bounded linear operator $U \in \mathscr{L}(H, K)$ between two Hilbert spaces is called a unitary operator if it is invertible and for every $x, y \in H$,

$$
\langle x, y\rangle_{H}=\langle U x, U y\rangle_{K} .
$$

Two Hilbert spaces $H$ and $K$ are unitarily equivalent if there exists a unitary operator $U \in \mathscr{L}(H, K)$.

Corollary D.33. Every infinite dimensional separable Hilbert space $H$ is unitarily equivalent to $l^{2}$.

Proof. Choose an orthonormal basis $\left(e_{n}\right)_{n \in \mathbb{N}}$ of $H$ (which exists by Corollary D.27), and define $U: H \rightarrow l^{2}$ by $U(x)=\left(\left\langle x, e_{n}\right\rangle\right)_{n \in \mathbb{N}}$. Then $\langle x, y\rangle_{H}=\langle U(x), U(y)\rangle_{l^{2}}$ by Theorem D. 31 ; in particular, $U$ is bounded, isometric and injective. The fact that $U$ is surjective, that is, that $\sum_{n} c_{n} e_{n}$ converges for every $c=\left(c_{n}\right) \in l^{2}$, follows as in the proof of Lemma D. 30 (i).

Clearly, if a sequence $\left(e_{n}\right)$ in a Hilbert space $H$ is an orthonormal basis, then necessarily $H$ is separable by Lemma D.21. Hence, the equivalent statements of Theorem D. 31 are only satisfied in separable Hilbert spaces. In most of the applications (if not all!), we will only deal with separable Hilbert spaces so that Theorem D. 31 is sufficient for our purposes.

However, what is true in general Hilbert spaces? The following sequence of results generalizes the preceeding results to arbitrary Hilbert spaces.

Definition D.34. Let $X$ be a normed space, $\left(x_{i}\right)_{i \in I}$ be a family. We say that the series $\sum_{i \in I} x_{i}$ converges unconditionally if the set $I_{0}:=\left\{i \in I: x_{i} \neq 0\right\}$ is countable, and for every bijective $\varphi: \mathbb{N} \rightarrow I_{0}$ the series $\sum_{n=1}^{\infty} x_{\varphi(n)}$ converges.

Corollary D. 35 (Bessel's inequality, general case). Let $H$ be an inner product space, $\left(e_{l}\right)_{l \in I} \subset H$ an orthonormal system. Then, for every $x \in H$, the set $\{l \in I$ : $\left.\left\langle x, e_{l}\right\rangle \neq 0\right\}$ is countable and

$$
\begin{equation*}
\sum_{l \in I}\left|\left\langle x, e_{l}\right\rangle\right|^{2} \leq\|x\|^{2} \tag{D.3}
\end{equation*}
$$

Proof. By Bessel's inequality, the sets $\left\{l \in I:\left|\left\langle x, e_{l}\right\rangle\right| \geq 1 / n\right\}$ must be finite for every $n \in \mathbb{N}$. The countability of $\left\{l \in I:\left\langle x, e_{l}\right\rangle \neq 0\right\}$ follows. The inequality (D.3) is then a direct consequence of Bessel's inequality.

Lemma D.36. Let $H$ be a Hilbert space, $\left(e_{l}\right)_{l \in I} \subset H$ an orthonormal system. Then:
a) For every $x \in H$, the series $\sum_{l \in I}\left\langle x, e_{l}\right\rangle e_{l}$ converges unconditionally.
b) $P: H \rightarrow H, x \mapsto \sum_{l \in I}\left\langle x, e_{l}\right\rangle e_{l}$ is the orthogonal projection onto $\overline{\operatorname{span}}\left\{e_{l}: l \in I\right\}$.

Corollary D.37. Every Hilbert space admits an orthonormal basis.
Proof. If $H$ is separable, the claim follows directly from the Gram-Schmidt process and has already been stated in Corollary D.27. In general, one may argue as follows:

The set of all orthonormal systems in $H$ forms a partially ordered set by inclusion. Given a totally ordered collection of orthonormal systems, the union of all vectors contained in all systems in this collection forms a supremum. By Zorn's lemma, there exists an orthonormal system $\left(e_{l}\right)_{l \in I}$ which is maximal. It follows from Bessel's inequality (D.3) that this system is actually an orthonormal basis.

Theorem D. 31 remains true for arbitrary Hilbert spaces when replacing the countable orthonormal system $\left(e_{n}\right)_{n \in \mathbb{N}}$ by an arbitrary orthonormal system $\left(e_{l}\right)_{l \in I}$.

## D. 3 * Fourier series

In the following we will identify the space $L^{1}(0,2 \pi)$ with

$$
L_{2 \pi}^{1}(\mathbb{R}):=\left\{f: \mathbb{R} \rightarrow \mathbb{C} \text { measurable, } 2 \pi \text {-periodic : } \int_{0}^{2 \pi}|f| d \lambda<\infty\right\}
$$

Similarly, we identify $L^{2}(0,2 \pi)$ with $L_{2 \pi}^{2}(\mathbb{R})$, and we define

$$
C_{2 \pi}(\mathbb{R}):=\{f \in C(\mathbb{R}): f \text { is } 2 \pi \text {-periodic }\}
$$

Definition D.38. For every $f \in L^{1}(0,2 \pi)=L_{2 \pi}^{1}(\mathbb{R})$ and every $n \in \mathbb{Z}$ we call

$$
\hat{f}(n):=\frac{1}{2 \pi} \int_{0}^{2 \pi} f(t) e^{-i n t} d t
$$

the $n$-th Fourier coefficient of $f$. The sequence $\hat{f}=(\hat{f}(n))$ is called the Fourier transform of $f$. The formal series $\frac{1}{\sqrt{2 \pi}} \sum_{n \in \mathbb{Z}} \hat{f}(n) e^{i n \cdot}$ is called the Fourier series of $f$.

Lemma D.39. For every $f \in L^{1}(0,2 \pi)=L_{2 \pi}^{1}(\mathbb{R})$ we have $\hat{f} \in l^{\infty}(\mathbb{Z})$ and the Fourier transform ${ }^{\wedge}: L^{1}(0,2 \pi) \rightarrow l^{\infty}$ is a bounded, linear operator. More precisely,

$$
\|\hat{f}\|_{\infty} \leq \frac{1}{2 \pi}\|f\|_{1}, \quad f \in L^{1}(0,2 \pi)
$$

Proof. For every $f \in L^{1}(0,2 \pi)$ and every $n \in \mathbb{Z}$,

$$
|\hat{f}(n)|=\frac{1}{2 \pi}\left|\int_{0}^{2 \pi} f(t) e^{-i n t} d t\right| \leq \frac{1}{2 \pi} \int_{0}^{2 \pi}|f(t)| d t
$$

This proves that $\hat{f} \in l^{\infty}$ and the required bound on $\|\hat{f}\|_{\infty}$. Linearity of ${ }^{\wedge}$ is clear.
Lemma D. 40 (Riemann-Lebesgue). For every $f \in L^{1}(0,2 \pi)=L_{2 \pi}^{1}(\mathbb{R})$ we have $\hat{f} \in c_{0}(\mathbb{Z})$, i.e.

$$
\lim _{|n| \rightarrow \infty}|\hat{f}(n)|=0
$$

Proof. Let $f \in L^{1}(0,2 \pi)=L_{2 \pi}^{1}(\mathbb{R})$ and $n \in \mathbb{Z}, n \neq 0$. Then

$$
\begin{aligned}
\hat{f}(n) & =\frac{1}{2 \pi} \int_{0}^{2 \pi} f(t) e^{-i n t} d t \\
& =\frac{1}{4 \pi} \int_{0}^{2 \pi} f(t) e^{-i n t}\left(1-e^{i \pi \frac{n}{n}}\right) d t \\
& =\frac{1}{4 \pi} \int_{0}^{2 \pi} f(t)\left(e^{-i n t}-e^{-i n\left(t-\frac{\pi}{n}\right)}\right) d t \\
& =\frac{1}{4 \pi} \int_{0}^{2 \pi}\left(f(t)-f\left(t+\frac{\pi}{n}\right)\right) e^{-i n t} d t
\end{aligned}
$$

so that

$$
|\hat{f}(n)| \leq \frac{1}{4 \pi} \int_{0}^{2 \pi}\left|f(t)-f\left(t+\frac{\pi}{n}\right)\right| d t
$$

Hence, if $f=1_{O} \in L^{1}(0,2 \pi)$ for some open set $O \subset[0,2 \pi]$, then $\hat{f} \in c_{0}(\mathbb{Z})$ by Lebesgue dominated convergence theorem. On the other hand, since span $\left\{1_{O}: O \subset\right.$ $[0,2 \pi]$ open $\}$ is dense in $L^{1}(0,2 \pi)$, since the Fourier transform is bounded with values in $l^{\infty}(\mathbb{Z})$ (Lemma D.39), and since $c_{0}(\mathbb{Z})$ is a closed subspace of $l^{\infty}(\mathbb{Z})$, we find that $\hat{f} \in c_{0}(\mathbb{Z})$ for every $f \in L^{1}(0,2 \pi)$.

Remark D.41. At the end of the proof of the Lemma of Riemann-Lebesgue, we used the following general principle: if $T \in \mathscr{L}(X, Y)$ is a bounded linear operator between two normed linear spaces $X, Y$, and if $M \subset X$ is dense, then $\operatorname{ran} T \subset \overline{T(M)}$. We used in addition that $c_{0}(\mathbb{Z})$ is closed in $l^{\infty}(\mathbb{Z})$.

Theorem D.42. Let $f \in C_{2 \pi}(\mathbb{R})$ be differentiable in some point $s \in \mathbb{R}$. Then

$$
f(s)=\sum_{n \in \mathbb{Z}} \hat{f}(n) e^{i n s}
$$

Proof. Note that for $f_{s}(t):=f(s+t)$,

$$
\hat{f}_{s}(n)=\frac{1}{2 \pi} \int_{0}^{2 \pi} f(s+t) e^{-i n t} d t=\frac{1}{2 \pi} \int_{0}^{2 \pi} f(t) e^{-i n(t-s)} d t=e^{i n s} \hat{f}(n)
$$

Hence, replacing $f$ by $f_{s}$, if necessary, we may without loss of generality assume that $s=0$. Moreover, replacing $f$ by $f-f(0)$, if necessary, we may without loss of generality assume that $f(0)=0$. We hence have to show that if $f$ is differentiable in 0 and if $f(0)=0$, then $\sum_{n \in \mathbb{Z}} \hat{f}(n)=0$.

Let $g(t):=\frac{f(t)}{1-e^{i t}}$. Since $f$ is differentiable in $0, f(0)=0$, and since $f$ is $2 \pi$ periodic, the function $g$ belongs to $C_{2 \pi}(\mathbb{R})$. By the Lemma of Riemann-Lebesgue, $\hat{g} \in c_{0}(\mathbb{Z})$. Note that

$$
\hat{f}(n)=\frac{1}{2 \pi} \int_{0}^{2 \pi} g(t)\left(1-e^{i t}\right) e^{-i n t} d t=\hat{g}(n)-\hat{g}(n-1)
$$

Hence,

$$
\begin{aligned}
\sum_{k=-n}^{n} \hat{f}(k) & =\sum_{k=-n}^{n} \hat{g}(k)-\hat{g}(k-1) \\
& =\hat{g}(n)-\hat{g}(-n-1) \rightarrow 0 \quad(n \rightarrow \infty)
\end{aligned}
$$

This is the claim.
Corollary D.43. For every $f \in C_{2 \pi}^{1}(\mathbb{R}):=C_{2 \pi}(\mathbb{R}) \cap C^{1}(\mathbb{R})$ and every $t \in \mathbb{R}$

$$
f(t)=\sum_{n \in \mathbb{Z}} \hat{f}(n) e^{i n t}
$$

Remark D.44. We will see that the convergence in the preceeding corollary is even uniform in $t \in \mathbb{R}$.

Throughout the following, we equip the space $L^{2}(0,2 \pi)=L_{2 \pi}^{2}(\mathbb{R})$ with the scalar product given by

$$
\langle f, g\rangle:=\frac{1}{2 \pi} \int_{0}^{2 \pi} f(t) \overline{g(t)} d t
$$

which differs from the usual scalar product by the factor $\frac{1}{2 \pi}$.
Lemma D.45. The space $C_{2 \pi}^{1}(\mathbb{R})$ is dense in $L_{2 \pi}^{2}(\mathbb{R})$.
Proof. We first prove that $C([0,2 \pi])$ is dense in $L^{2}(0,2 \pi)=L_{2 \pi}^{2}(\mathbb{R})$. For this, consider first a characteristic function $f=1_{(a, b)} \in L^{2}(0,2 \pi)$. Let $\left(g_{n}\right) \subset C([0,2 \pi])$ be defined by

$$
g_{n}(t):= \begin{cases}1, & t \in[a, b] \\ 1+n(t-a), & t \in[a-1 / n, a), \\ 1-n(t-b), & t \in(b, b+1 / n] \\ 0, & \text { else. }\end{cases}
$$

It is then easy to see that $\lim _{n \rightarrow \infty}\left\|f-g_{n}\right\|_{L^{2}}=0$, so that $f=1_{(a, b)} \in \overline{C([0,2 \pi])}\|\cdot\|_{L^{2}}$.
In the second step, consider a characteristic function $f=1_{A}$ of an arbitrary Borel set $A \in \mathscr{B}([0,2 \pi])$, and let $\varepsilon>0$. By outer regularity of the Lebesgue measure, there exists an open set $O \supset A$ such that $\lambda(O \backslash A)<\varepsilon^{2}$. Recall that $O$ is the countable union of mutually disjoint intervals. Since $O$ has finite measure, there exist finitely many (mutually disjoint) intervals $\left(a_{n}, b_{n}\right) \subset O(1 \leq n \leq N)$ such that $\lambda\left(O \backslash \bigcup_{n=1}^{N}\left(a_{n}, b_{n}\right)\right) \leq \varepsilon^{2}$. By the preceeding step, for every $1 \leq n \leq N$ there exists $g_{n} \in C([0,2 \pi])$ such that $\left\|1_{\left(a_{n}, b_{n}\right)}-g_{n}\right\|_{2} \leq \frac{\varepsilon}{N}$. Let $g:=\sum_{n=1}^{N} g_{n} \in C([0,2 \pi])$. Then

$$
\begin{aligned}
\|f-g\|_{2} & \leq\left\|1_{A}-1_{O}\right\|_{2}+\left\|1_{O}-1_{\cup_{n=1}^{N}\left(a_{n}, b_{n}\right)}\right\|_{2}+\left\|1_{\cup_{n=1}^{N}\left(a_{n}, b_{n}\right)}-g\right\|_{2} \\
& \leq \varepsilon+\varepsilon+\left\|\sum_{n=1}^{N}\left(1_{\left(a_{n}, b_{n}\right)}-g_{n}\right)\right\|_{2}
\end{aligned}
$$

$$
\leq 3 \varepsilon
$$

 $A \in \mathscr{B}([0,2 \pi])\}=L^{2}(0,2 \pi)$, we find that $C([0,2 \pi])$ is dense in $L^{2}(0,2 \pi)$.

It remains to show that $C_{2 \pi}^{1}(\mathbb{R})$ is dense in $C([0,2 \pi])$ for the norm $\|\cdot\|_{2}$. So let $f \in C([0,2 \pi])$ and let $\varepsilon>0$. By Weierstrass' theorem, there exists a function $g_{0} \in C^{\infty}([0,2 \pi])$ (even a polynomial!) such that $\left\|f-g_{0}\right\|_{\infty} \leq \varepsilon$. Let $g_{1} \in C^{1}([0,2 \pi])$ be such that $g_{1}(2 \pi)=g_{1}^{\prime}(2 \pi)=0, g_{1}(0)=g_{0}(2 \pi)-g_{0}(0)$ and $g_{1}^{\prime}(0)=g_{0}^{\prime}(2 \pi)-$ $g_{0}^{\prime}(0)$ and $\left\|g_{1}\right\|_{2} \leq \varepsilon$. Such a function $g_{1}$ exists: it suffices for example to consider functions for which the derivative is of the form

$$
g_{1}^{\prime}(t)= \begin{cases}g_{0}(2 \pi)-g_{0}(0)+c t, & t \in\left[0, h_{1}\right], \\ g_{0}(2 \pi)-g_{0}(0)+c h_{1}+d\left(t-h_{1}\right), & t \in\left(h_{1}, h_{2}\right), \\ 0, & t \in\left[h_{2}, 2 \pi\right]\end{cases}
$$

with appropriate constants $0 \leq h_{1} \leq h_{2}$ and $c, d \in \mathbb{C}$. Having chosen $g_{1}$, we let $g=g_{0}+g_{1}$ and we calculate that

$$
\|f-g\|_{2} \leq\left\|f-g_{0}\right\|_{2}+\left\|g_{1}\right\|_{2} \leq 2 \varepsilon .
$$

Since $g$ extends to a function in $C_{2 \pi}^{1}(\mathbb{R})$, we have thus proved that $C_{2 \pi}^{1}(\mathbb{R})$ is dense in $L_{2 \pi}^{2}(\mathbb{R})$.
Remark D.46. An adaptation of the above proof actually shows that for every $1 \leq$ $p<\infty$ and every compact interval $[a, b] \subset \mathbb{R}$, the space $C([a, b])$ is dense in $L^{p}(a, b)$. A further application of Weierstrass' theorem actually shows that the space of all polynomials is dense in $L^{p}(a, b)$. In particular, we may obtain the following result.
Corollary D.47. The space $L^{p}(a, b)$ is separable if $1 \leq p<\infty$. The space $L^{\infty}(a, b)$ is not separable.
Corollary D.48. Let $e_{n}(t):=e^{\text {int }}, n \in \mathbb{Z}, t \in \mathbb{R}$. Then $\left(e_{n}\right)_{n \in \mathbb{Z}}$ is an orthonormal basis in $L_{2 \pi}^{2}(\mathbb{R})$.
Proof. The fact that $\left(e_{n}\right)_{n \in \mathbb{Z}}$ is an orthonormal system in $L_{2 \pi}^{2}(\mathbb{R})$ is an easy calculation. We only have to prove that span $\left\{e_{n}: n \in \mathbb{Z}\right\}$ is dense in $L_{2 \pi}^{2}(\mathbb{R})$. Note that $\hat{f}(n)=\left(f, e_{n}\right)$ for every $f \in L_{2 \pi}^{2}(\mathbb{R})$ and every $n \in \mathbb{Z}$. By Lemma D.30, we know that for every $f \in L_{2 \pi}^{2}(\mathbb{R})$

$$
g:=\sum_{n \in \mathbb{Z}} \hat{f}(n) e_{n} \text { exists in } L_{2 \pi}^{2}(\mathbb{R})
$$

In particular, a subsequence of $\left(\sum_{n=-k}^{k} \hat{f}(n) e_{n}\right)$ converges almost everywhere to $g$. But by Corollary D. 43 we know that $\left(\sum_{n=-k}^{k} \hat{f}(n) e_{n}\right)$ converges pointwise everywhere to $f$ if $f \in C_{2 \pi}^{1}(\mathbb{R})$. As a consequence, for every $f \in C_{2 \pi}^{1}(\mathbb{R})$,

$$
\lim _{k \rightarrow \infty} \sum_{n=-k}^{k} \hat{f}(n) e_{n}=f \text { in } L_{2 \pi}^{2}(\mathbb{R})
$$

so that $\operatorname{span}\left\{e_{n}: n \in \mathbb{Z}\right\}$ is dense in $\left(C_{2 \pi}^{1}(\mathbb{R}),\|\cdot\|_{L_{2 \pi}^{2}}\right)$. Since $C_{2 \pi}^{1}(\mathbb{R})$ is dense in $L_{2 \pi}^{2}(\mathbb{R})$ by Lemma D.45, we find that $\left(e_{n}\right)_{n \in \mathbb{Z}}$ is an orthonormal basis in $L_{2 \pi}^{2}(\mathbb{R})$.
Theorem D. 49 (Plancherel). For every $f \in L_{2 \pi}^{2}(\mathbb{R})$ we have $\hat{f} \in l^{2}(\mathbb{Z})$ and the Fourier transform ${ }^{\wedge}: L_{2 \pi}^{2}(\mathbb{R}) \rightarrow l^{2}(\mathbb{Z})$ is an isometric isomorphism. Moreover, for every $f \in L_{2 \pi}^{2}(\mathbb{R})$,

$$
\sum_{n \in \mathbb{Z}} \hat{f}(n) e_{n}=f \text { in } L_{2 \pi}^{2}(\mathbb{R})
$$

that is, the Fourier series of $f$ converges to $f$ in the $L^{2}$ sense.

Proof. By Corollary D.48, the sequence $\left(e_{n}\right)_{n \in \mathbb{Z}}$ is an orthonormal basis in $L_{2 \pi}^{2}(\mathbb{R})$. Moreover, recall that for every $f \in L_{2 \pi}^{2}(\mathbb{R})$ and every $n \in \mathbb{Z}, \hat{f}(n)=\left\langle f, e_{n}\right\rangle$. Hence, by Theorem D.31, $\hat{f} \in l^{2}(\mathbb{Z}), f=\sum_{n \in \mathbb{Z}} \hat{f}(n) e_{n}$, and $\|f\|_{L_{2 \pi}^{2}}=\|\hat{f}\|_{l^{2}}$ (the last property being Parseval's identity).

Corollary D.50. Let $f \in C_{2 \pi}(\mathbb{R})$ be such that $\hat{f} \in l^{1}(\mathbb{Z})$. Then

$$
\sum_{n \in \mathbb{Z}} \hat{f}(n) e_{n}=f \text { in } C_{2 \pi}(\mathbb{R})
$$

that is, the Fourier series of $f$ converges uniformly to $f$.
Proof. Note that for every $n \in \mathbb{Z},\left\|e_{n}\right\|_{\infty}=1$. The assumption $\hat{f} \in l^{1}(\mathbb{Z})$ therefore implies that the series $\sum_{n \in \mathbb{Z}} \hat{f}(n) e_{n}$ converges absolutely in $C_{2 \pi}(\mathbb{R})$, i.e. for the uniform norm $\|\cdot\|_{\infty}$. Since $\left(C_{2 \pi}(\mathbb{R}),\|\cdot\|_{\infty}\right)$ is complete, the series $\sum_{n \in \mathbb{Z}} \hat{f}(n) e_{n}$ converges uniformly to some element $g \in C_{2 \pi}(\mathbb{R})$. By Plancherel, $g=f$.

Remark D.51. The assumption $\hat{f} \in l^{1}(\mathbb{Z})$ in Corollary D. 50 is essential. For general $f \in C_{2 \pi}(\mathbb{R})$, the Fourier series $\sum_{n \in \mathbb{Z}} \hat{f}(n) e_{n}$ need not not converge uniformly. Questions regarding the convergence of Fourier series (which type of convergence? for which function?) can go deeply into the theory of harmonic analysis and answers are sometimes quite involved. The $L^{2}$ theory gives in this context satisfactory answers with relatively easy proofs (see Plancherel's theorem). For continuous functions we state the following result without giving a proof.

Theorem D. 52 (Féjer). For every $f \in C_{2 \pi}(\mathbb{R})$ one has

$$
\lim _{K \rightarrow \infty} \frac{1}{K} \sum_{k=1}^{K} \sum_{n=-k}^{k} \hat{f}(n) e_{n}=f \text { in } C_{2 \pi}(\mathbb{R})
$$

i.e. the Fourier series of $f$ converges in the Césaro mean uniformly to $f$.

## D. 4 Linear functionals on Hilbert spaces

In this section, we discuss bounded functionals on Hilbert spaces. Compared to the case of bounded linear functionals on general Banach spaces, the case of bounded linear functionals on Hilbert spaces is considerably easy but it has far reaching consequences.

Theorem D. 53 (Riesz-Fréchet). Let $H$ be a Hilbert space. Then for every bounded linear functional $\varphi \in H^{\prime}$ there exists a unique $y \in H$ such that

$$
\varphi(x)=\langle x, y\rangle \quad \text { for every } x \in H
$$

Proof. Uniqueness. Let $y_{1}, y_{2} \in H$ be two elements such that

$$
\varphi(x)=\left\langle x, y_{1}\right\rangle=\left\langle x, y_{2}\right\rangle \quad \text { for every } x \in H
$$

Then $\left\langle x, y_{1}-y_{2}\right\rangle=0$ for every $x \in H$, in particular also for $x=y_{1}-y_{2}$. This implies $\left\|y_{1}-y_{2}\right\|^{2}=0$, that is, $y_{1}=y_{2}$.

Existence. We may assume that $\varphi \neq 0$ since the case $\varphi=0$ is trivial. Let $\tilde{y} \in$ $(\operatorname{ker} \varphi)^{\perp} \backslash\{0\}$. Since $H \neq \operatorname{ker} \varphi$ and since $\operatorname{ker} \varphi$ is closed, such a $\tilde{y}$ exists. Next, let

$$
y:=\overline{\varphi(\tilde{y})} /\|\tilde{y}\|^{2} \tilde{y} .
$$

Note that $\varphi(y)=\|y\|^{2}=\langle y, y\rangle$. Recall that every $x \in H$ can be uniquely written as $x=x_{0}+\lambda y$ with $x_{0} \in \operatorname{ker} \varphi$ and $\lambda \in \mathbb{K}$ so that $\lambda y \in(\operatorname{ker} \varphi)^{\perp}$. Note that $(\operatorname{ker} \varphi)^{\perp}$ is one-dimensional. Hence, for every $x \in H$,

$$
\begin{aligned}
\varphi(x) & =\varphi\left(x_{0}+\lambda y\right) \\
& =\varphi\left(x_{0}\right)+\lambda \varphi(y) \\
& =\lambda \varphi(y) \\
& =\lambda\langle y, y\rangle \\
& =\langle\lambda y, y\rangle \\
& =\left\langle x_{0}, y\right\rangle+\langle\lambda y, y\rangle \\
& =\langle x, y\rangle .
\end{aligned}
$$

The claim is proved.
Corollary D.54. Let $J: H \rightarrow H^{\prime}$ be the mapping which maps to every $y \in H$ the functional Jy $\in H^{\prime}$ given by $J y(x)=\langle x, y\rangle$. Then $J$ is antilinear if $\mathbb{K}=\mathbb{C}$ and linear if $\mathbb{K}=\mathbb{R}$. Moreover, $J$ is isometric and bijective.

Proof. The fact that $J$ is isometric follows from the Cauchy-Schwarz inequality. Antilinearity (or linearity in case $\mathbb{K}=\mathbb{R}$ ) follows from the sesquilinearity (resp. bilinearity) of the scalar product on $H$. Since $J$ is isometric, it is injective. The surjectivity of $J$ follows from Theorem D.53.

Remark D.55. The theorem of Riesz-Fréchet allows us to identify any (real) Hilbert space $H$ with its dual space $H^{\prime}$. Note, however, that there are situations in which one does not identify $H^{\prime}$ with $H$. This is for example the case when $V$ is a second Hilbert space which embeds continuously and densely into $H$, that is, for which there exists a bounded, injective $J: V \rightarrow H$ with dense range.

## D. 5 Weak convergence in Hilbert spaces

Definition D.56. Let $H$ be a Hilbert space. We say that a sequence $\left(x_{n}\right) \subset H$ converges weakly to some element $x \in H$ if for every $y \in H$ one has $\lim _{n \rightarrow \infty}\left\langle x_{n}, y\right\rangle=$ $\langle x, y\rangle$. We write $x_{n} \rightharpoonup x$ or $x_{n} \xrightarrow{\text { weak }} x$ if $\left(x_{n}\right)$ converges weakly to $x$.

Theorem D.57. Every bounded sequence $\left(x_{n}\right)$ in a Hilbert space $H$ admits a weakly convergent subsequence, that is, there exists $x \in H$ and there exists a subsequence $\left(x_{n_{k}}\right)$ of $\left(x_{n}\right)$ such that $x_{n_{k}} \xrightarrow{\text { weak }} x$.

In the proof of this theorem, we will use the following general result.
Lemma D.58. Let $X$ and $Y$ be two normed spaces, let $\left(T_{n}\right) \in \mathscr{L}(X, Y)$ be a bounded sequence of bounded operators. Assume that there exists a dense set $M \subset X$ such that $\lim _{n \rightarrow \infty} T_{n} x$ exists for every $x \in M$. Then $\lim _{n \rightarrow \infty} T_{n} x=: T x$ exists for every $x \in X$ and $T \in \mathscr{L}(X, Y)$.
Proof. Define $T x:=\lim _{n \rightarrow \infty} T_{n} x$ for every $x \in \operatorname{span} M$. Then

$$
\|T x\|=\lim _{n \rightarrow \infty}\left\|T_{n} x\right\| \leq \sup _{n}\left\|T_{n}\right\|\|x\|,
$$

that is. $T: \operatorname{span} M \rightarrow Y$ is a bounded linear operator. Since $M$ is dense in $X, T$ admits a unique bounded extension $T: X \rightarrow Y$.

Let $x \in X$ and $\varepsilon>0$. Since $M$ is dense in $X$, there exists $y \in M$ such that $\|x-y\| \leq$ $\varepsilon$. By assumption, there exists $n_{0}$ such that for every $n \geq n_{0}$ we have $\left\|T_{n} y-T y\right\| \leq \varepsilon$. Hence, for every $n \geq n_{0}$,

$$
\begin{aligned}
\left\|T_{n} x-T x\right\| & \leq\left\|T_{n} x-T_{n} y\right\|+\left\|T_{n} y-T y\right\|+\|T y-T x\| \\
& \leq \sup _{n}\left\|T_{n}\right\|\|x-y\|+\varepsilon+\|T\|\|x-y\| \\
& \leq \varepsilon\left(\sup _{n}\left\|T_{n}\right\|+1+\|T\|\right),
\end{aligned}
$$

and therefore $\lim _{n \rightarrow \infty} T_{n} x=T x$.
Proof (Proof of Theorem D.57). As in the proof of the Arzela-Ascoli theorem (Theorem B.43), we use Cantor's diagonal sequence argument. Let $\left(x_{n}\right)$ be a bounded sequence in $H$. We first assume that $H$ is separable, and we let $\left(y_{m}\right) \subset H$ be a dense sequence.

Since $\left(\left\langle x_{n}, y_{1}\right\rangle\right)$ is bounded by the boundedness of $\left(x_{n}\right)$, there exists a subsequence $\left(x_{\varphi_{1}(n)}\right)$ of $\left(x_{n}\right)\left(\varphi_{1}: \mathbb{N} \rightarrow \mathbb{N}\right.$ is increasing, unbounded) such that

$$
\lim _{n \rightarrow \infty}\left\langle x_{\varphi_{1}(n)}, y_{1}\right\rangle \text { exists. }
$$

Similarly, there exists a subsequence $\left(x_{\varphi_{2}(n)}\right)$ of $\left(x_{\varphi_{1}(n)}\right)$ such that

$$
\lim _{n \rightarrow \infty}\left\langle x_{\varphi_{2}(n)}, y_{2}\right\rangle \text { exists. }
$$

Note that for this subsequence, we also have that

$$
\lim _{n \rightarrow \infty}\left\langle x_{\varphi_{2}(n)}, y_{1}\right\rangle \text { exists. }
$$

Iterating this argument, we find a subsequence $\left(x_{\varphi_{3}(n)}\right)$ of $\left(x_{\varphi_{2}(n)}\right)$ and finally for every $m \in \mathbb{N}, m \geq 2$, a subsequence $\left(x_{\varphi_{m}(n)}\right)$ of $\left(x_{\varphi_{m-1}(n)}\right)$ such that

$$
\lim _{n \rightarrow \infty}\left\langle x_{\varphi_{m}(n)}, y_{j}\right\rangle \text { exists for every } 1 \leq j \leq m
$$

Let $\left(x_{n}^{\prime}\right):=\left(x_{\varphi_{n}(n)}\right)$ be the 'diagonal sequence'. Then $\left(x_{n}^{\prime}\right)$ is a subsequence of $\left(x_{n}\right)$ and

$$
\lim _{n \rightarrow \infty}\left\langle x_{n}^{\prime}, y_{m}\right\rangle \text { exists for every } m \in \mathbb{N}
$$

By Lemma D. 58 and the Riesz-Fréchet representation theorem (Theorem D.53), there exists $x \in H$ such that

$$
\lim _{n \rightarrow \infty}\left\langle x_{n}^{\prime}, y\right\rangle=\langle x, y\rangle \text { for every } y \in H
$$

and the claim is proved in the case when $H$ is separable.
If $H$ is not separable as we first assumed, then one may replace $H$ by $\tilde{H}:=$ $\overline{\operatorname{span}}\left\{x_{n}: n \in \mathbb{N}\right\}$ which is separable. By the above, there exists $x \in \tilde{H}$ and a subsequence of $\left(x_{n}\right)$ (which we denote again by $\left(x_{n}\right)$ ) such that for every $y \in \tilde{H}$,

$$
\lim _{n \rightarrow \infty}\left\langle x_{n}, y\right\rangle=\langle x, y\rangle,
$$

that is, $\left(x_{n}\right)$ converges weakly in $\tilde{H}$. On the other hand, for every $y \in \tilde{H}^{\perp}$ and every $n$,

$$
\left\langle x_{n}, y\right\rangle=\langle x, y\rangle=0
$$

The decomposition $H=\tilde{H} \oplus \tilde{H}^{\perp}$ therefore yields that $\left(x_{n}\right)$ converges weakly in $H$.

## Appendix E

## Dual spaces and weak convergence

## E. 1 The theorem of Hahn-Banach

Given a normed space $X$, we denote by $X^{\prime}:=\mathscr{L}(X, \mathbb{K})$ the space of all bounded linear functionals on $X$. Recall that $X^{\prime}$ is always a Banach space by Corollary B. 35 of Chapter B.

However, a priori it is not clear whether there exists any bounded linear functional on a normed space $X$ (apart from the zero functional). This fundamental question and the analysis of dual spaces (analysis of functionals) shall be developed in this chapter.

The existence of nontrivial bounded functionals is guaranteed by the HahnBanach theorem which actually admits several versions. However, before stating the first version, we need the following definition.

Definition E.1. Let $X$ be a real or complex vector space. A function $p: X \rightarrow \mathbb{R}$ is called sublinear if
(i) $p(\lambda x)=\lambda p(x)$ for every $\lambda>0, x \in X$, and
(ii) $p(x+y) \leq p(x)+p(y)$ for every $x, y \in X$.

Example E.2. On a normed space $X$, the norm $\|\cdot\|$ is sublinear. Every linear $p$ : $X \rightarrow \mathbb{R}$ is sublinear.

Theorem E. 3 (Hahn-Banach; version of linear algebra, real case). Let $X$ be a real vector space, $U \subset X$ a linear subspace, and $p: X \rightarrow \mathbb{R}$ sublinear. Let $\varphi: U \rightarrow \mathbb{R}$ be linear such that

$$
\varphi(x) \leq p(x) \text { for all } x \in U
$$

Then there exists a linear $\tilde{\varphi}: X \rightarrow \mathbb{R}$ such that $\tilde{\varphi}(x)=\varphi(x)$ for every $x \in U$ (that is, $\tilde{\varphi}$ is an extension of $\varphi$ ) and

$$
\begin{equation*}
\tilde{\varphi}(x) \leq p(x) \text { for all } x \in X \tag{E.1}
\end{equation*}
$$

The following lemma asserts that this version of Hahn-Banach is true in the special case when $X / U$ has dimension 1. It is an essential step in the proof of Theorem E.3.

Lemma E.4. Take the assumptions of Theorem E. 3 and assume in addition that $\operatorname{dim} X / U=1$. Then the assertion of Theorem E. 3 is true.

Proof. If $\operatorname{dim} X / U=1$, then there exists $x_{0} \in X \backslash U$ such that every $x \in X$ can be uniquely written in the form $x=u+\lambda x_{0}$ with $u \in U$ and $\lambda \in \mathbb{R}$. So we define $\tilde{\varphi}: X \rightarrow \mathbb{R}$ by

$$
\tilde{\varphi}(x):=\tilde{\varphi}\left(u+\lambda x_{0}\right):=\varphi(u)+\lambda r
$$

where $r \in \mathbb{R}$ is a parameter which has to be chosen such that (E.1) holds, that is, such that for every $u \in U, \lambda \in \mathbb{R}$,

$$
\begin{equation*}
\varphi(u)+\lambda r \leq p\left(u+\lambda x_{0}\right) \tag{E.2}
\end{equation*}
$$

If $\lambda=0$, then this condition clearly holds for every $u \in U$ by the assumption on $\varphi$. If $\lambda>0$, then (E.2) holds for every $u \in U$ if and only if

$$
\begin{aligned}
& \lambda r \leq p\left(u+\lambda x_{0}\right)-\varphi(u) \text { for every } u \in U \\
\Leftrightarrow & r \leq p\left(\frac{u}{\lambda}+x_{0}\right)-\varphi\left(\frac{u}{\lambda}\right) \text { for every } u \in U \\
\Leftrightarrow & r \leq \inf _{v \in U} p\left(v+x_{0}\right)-\varphi(v) .
\end{aligned}
$$

Similarly, if $\lambda<0$, then (E.2) holds for every $u \in U$ if and only if

$$
\begin{aligned}
& \lambda r \leq p\left(u+\lambda x_{0}\right)-\varphi(u) \text { for every } u \in U \\
\Leftrightarrow & -r \leq p\left(\frac{u}{-\lambda}-x_{0}\right)-\varphi\left(\frac{u}{-\lambda}\right) \text { for every } u \in U \\
\Leftrightarrow & r \geq \sup _{w \in U} \varphi(w)-p\left(w-x_{0}\right)
\end{aligned}
$$

So it is possible to find an appropriate $r \in \mathbb{R}$ in the definition of $\tilde{\varphi}$ if and only if

$$
\varphi(w)-p\left(w-x_{0}\right) \leq p\left(v+x_{0}\right)-\varphi(v) \text { for all } v, w \in U
$$

or, equivalently, if

$$
\varphi(w)+\varphi(v) \leq p\left(v+x_{0}\right)+p\left(w-x_{0}\right) \text { for all } v, w \in U
$$

However, by the assumptions on $\varphi$ and $p$, for every $v, w \in U$,

$$
\varphi(w)+\varphi(v)=\varphi(w+v) \leq p(w+v)=p\left(v+x_{0}+w-x_{0}\right) \leq p\left(v+x_{0}\right)+p\left(w-x_{0}\right) .
$$

For the second step in the proof of Theorem E.3, we need the Lemma of Zorn.
Lemma E. 5 (Zorn). Let $(M, \leq)$ be a ordered set. Assume that every totally ordered subset $T \subset M$ (i.e. for every $x, y \in T$ one either has $x \leq y$ or $y \leq x$ ) admits an
upper bound. Then for every $x \in M$ there exists a maximal element $m \geq x$ (that is, an element $m$ such that $m \leq \tilde{m}$ implies $m=\tilde{m}$ for every $\tilde{m} \in M$ ).
Proof (Proof of Theorem E.3). Define the following set

$$
\begin{aligned}
M:=\left\{\left(V, \varphi_{V}\right):\right. & V \subset X \text { linear subspace }, U \subset V, \varphi_{V}: V \rightarrow \mathbb{R} \text { linear, s.t. } \\
& \left.\varphi(x)=\varphi_{V}(x)(x \in U) \text { and } \varphi_{V}(x) \leq p(x)(x \in V)\right\}
\end{aligned}
$$

and equip it with the order relation $\leq$ defined by

$$
\left(V_{1}, \varphi_{V_{1}}\right) \leq\left(V_{2}, \varphi_{V_{2}}\right): \Leftrightarrow V_{1} \subset V_{2} \text { and } \varphi_{V_{1}}(x)=\varphi_{V_{2}}(x) \text { for all } x \in V_{1}
$$

Then $(M, \leq)$ is an ordered set. Let $T=\left(\left(V_{i}, \varphi_{V_{i}}\right)\right)_{i \in I} \subset M$ be a totally ordered subset. Then the element $\left(V, \varphi_{V}\right) \in M$ defined by

$$
V:=\bigcup_{i \in I} V_{i} \text { and } \varphi_{V}(x)=\varphi_{V_{i}}(x) \text { for } x \in V_{i}
$$

is an upper bound of $T$. By the Lemma of Zorn, the set $M$ admits a maximal element $\left(X_{0}, \varphi_{X_{0}}\right)$. Assume that $X_{0} \neq X$. Then, by Lemma E.4, we could construct an element which is strictly larger than $\left(X_{0}, \varphi_{X_{0}}\right)$, a contradiction to the maximality of $\left(X_{0}, \varphi_{X_{0}}\right)$. Hence, $X=X_{0}$, and $\tilde{\varphi}:=\varphi_{X_{0}}$ is an element we are looking for.

The complex version of the Hahn-Banach theorem reads as follows.
Theorem E. 6 (Hahn-Banach; version of linear algebra, complex case). Let X be a complex vector space, $U \subset X$ a linear subspace, and $p: X \rightarrow \mathbb{R}$ sublinear. Let $\varphi: U \rightarrow \mathbb{C}$ be linear such that

$$
\operatorname{Re} \varphi(x) \leq p(x) \text { for all } x \in U
$$

Then there exists a linear $\tilde{\varphi}: X \rightarrow \mathbb{C}$ such that $\tilde{\varphi}(x)=\varphi(x)$ for every $x \in U$ (that is $\tilde{\varphi}$ is an extension of $\varphi$ ) and

$$
\begin{equation*}
\operatorname{Re} \tilde{\varphi}(x) \leq p(x) \text { for all } x \in X \tag{E.3}
\end{equation*}
$$

Proof. We may consider $X$ also as a real vector space. Note that $\psi(x):=\operatorname{Re} \varphi(x)$ is an $\mathbb{R}$-linear functional on $X$. By Theorem E.3, there exists an extension $\tilde{\psi}: X \rightarrow \mathbb{R}$ of $\psi$ satisfying

$$
\tilde{\psi}(x) \leq p(x) \text { for every } x \in X
$$

Let

$$
\tilde{\varphi}(x):=\tilde{\psi}(x)-i \tilde{\psi}(i x), \quad x \in X
$$

It is an exercise to show that $\tilde{\varphi}$ is $\mathbb{C}$-linear, that $\varphi(x)=\tilde{\varphi}(x)$ for every $x \in U$ and it is clear from the definition that $\operatorname{Re} \tilde{\varphi}(x)=\tilde{\psi}(x)$. Thus, $\tilde{\varphi}$ is a possible element we are looking for.

Theorem E. 7 (Hahn-Banach; extension of bounded linear functionals). Let $X$ be a normed space and $U \subset X$ a linear subspace. Then for every bounded linear
$u^{\prime}: U \rightarrow \mathbb{K}$ there exists a bounded linear extension $x^{\prime}: X \rightarrow \mathbb{K}\left(\right.$ that is, $\left.\left.x^{\prime}\right|_{U}=u^{\prime}\right)$ such that $\left\|x^{\prime}\right\|=\left\|u^{\prime}\right\|$.

Proof. We first assume that $X$ is a real normed space. The function $p: X \rightarrow \mathbb{R}$ defined by $p(x):=\left\|u^{\prime}\right\|\|x\|$ is sublinear and

$$
u^{\prime}(x) \leq p(x) \text { for every } x \in U
$$

By the first Hahn-Banach theorem (Theorem E.3), there exists a linear $x^{\prime}: X \rightarrow \mathbb{R}$ extending $u^{\prime}$ such that

$$
x^{\prime}(x) \leq p(x)=\left\|u^{\prime}\right\|\|x\| \text { for every } x \in X
$$

Replacing $x$ by $-x$, this implies that

$$
\left|x^{\prime}(x)\right| \leq\left\|u^{\prime}\right\|\|x\| \text { for every } x \in X
$$

Hence, $x^{\prime}$ is bounded and $\left\|x^{\prime}\right\| \leq\left\|u^{\prime}\right\|$. On the other hand, one trivially has

$$
\left\|x^{\prime}\right\|=\sup _{\substack{x \in X \\\|x\| \leq 1}}\left|x^{\prime}(x)\right| \geq \sup _{\substack{x \in U \\\|x\| \leq 1}}\left|x^{\prime}(x)\right|=\sup _{\substack{x \in U \\\|x\| \leq \leq 1}}\left|u^{\prime}(x)\right|=\left\|u^{\prime}\right\| .
$$

If $X$ is a complex normed space, then the second Hahn-Banach theorem (Theorem E.6) implies that there exists a linear $x^{\prime}: X \rightarrow \mathbb{C}$ such that

$$
\operatorname{Re} x^{\prime}(x) \leq p(x)=\left\|u^{\prime}\right\|\|x\| \text { for every } x \in X
$$

In particular,

$$
\left|x^{\prime}(x)\right|=\sup _{\theta \in[0,2 \pi]} \operatorname{Re} x^{\prime}\left(e^{i \theta} x\right) \leq\left\|u^{\prime}\right\|\|x\| \text { for every } x \in X
$$

so that again $x^{\prime}$ is bounded and $\left\|x^{\prime}\right\| \leq\left\|u^{\prime}\right\|$. The inequality $\left\|x^{\prime}\right\| \geq\left\|u^{\prime}\right\|$ follows as above.

Corollary E.8. If $X$ is a normed space, then for every $x \in X \backslash\{0\}$ there exists $x^{\prime} \in X^{\prime}$ such that

$$
\left\|x^{\prime}\right\|=1 \text { and } x^{\prime}(x)=\|x\| .
$$

In particular, $X^{\prime}$ separates the points of $X$, i.e. for every $x_{1}, x_{2} \in X, x_{1} \neq x_{2}$, there exists $x^{\prime} \in X^{\prime}$ such that $x^{\prime}\left(x_{1}\right) \neq x^{\prime}\left(x_{2}\right)$.

Proof. By the Hahn-Banach theorem (Theorem E.7), there exists an extension $x^{\prime} \in$ $X^{\prime}$ of the functional $u^{\prime}: \operatorname{span}\{x\} \rightarrow \mathbb{K}$ defined by $u^{\prime}(\lambda x)=\lambda\|x\|$ such that $\left\|x^{\prime}\right\|=$ $\left\|u^{\prime}\right\|=1$.

For the proof of the second assertion, set $x:=x_{1}-x_{2}$.
Corollary E.9. If $X$ is a normed space, then for every $x \in X$

$$
\begin{equation*}
\|x\|=\sup _{\substack{x^{\prime} \in X^{\prime} \\\left\|x^{\prime}\right\| \leq 1}}\left|x^{\prime}(x)\right| . \tag{E.4}
\end{equation*}
$$

Proof. For every $x^{\prime} \in X^{\prime}$ with $\left\|x^{\prime}\right\| \leq 1$ one has

$$
\left|x^{\prime}(x)\right| \leq\left\|x^{\prime}\right\|\|x\| \leq\|x\|
$$

which proves one of the required inequalities. The other inequality follows from Corollary E.8.

Remark E.10. The equality (E.4) should be compared to the definition of the norm of an element $x^{\prime} \in X^{\prime}$ :

$$
\left\|x^{\prime}\right\|=\sup _{\substack{x \in X \\\|x\| \leq 1}}\left|x^{\prime}(x)\right| .
$$

From now on, it will be convenient to use the following notation. Given a normed space $X$ and elements $x \in X, x^{\prime} \in X^{\prime}$, we write

$$
\left\langle x^{\prime}, x\right\rangle:=\left\langle x^{\prime}, x\right\rangle_{X^{\prime} \times X}:=x^{\prime}(x) .
$$

For the bracket $\langle\cdot, \cdot\rangle$, we note the following properties. The function

$$
\begin{aligned}
\langle\cdot, \cdot\rangle: X^{\prime} \times X & \rightarrow \mathbb{K}, \\
\left(x^{\prime}, x\right) & \mapsto\left\langle x^{\prime}, x\right\rangle=x^{\prime}(x)
\end{aligned}
$$

is bilinear and for every $x^{\prime} \in X^{\prime}, x \in X$,

$$
\left|\left\langle x^{\prime}, x\right\rangle\right| \leq\left\|x^{\prime}\right\|\|x\|
$$

The bracket $\langle\cdot, \cdot\rangle$ thus appeals to the notion of the scalar product on inner product spaces, and the last inequality appeals to the Cauchy-Schwarz inequality, but note, however, that the bracket is not a scalar product since it is defined on a pair of two different spaces. Moreover, even if $X=H$ is a complex Hilbert space, then the bracket differs from the scalar product in that it is bilinear instead of sesquilinear.

Corollary E.11. Let $X$ be a normed space, $U \subset X$ a closed linear subspace and $x \in X \backslash U$. Then there exists $x^{\prime} \in X^{\prime}$ such that

$$
x^{\prime}(x) \neq 0 \text { and } x^{\prime}(u)=0 \text { for every } u \in U
$$

Proof. Let $\pi: X \rightarrow X / U$ be the quotient map $(\pi(x)=x+U)$. Since $x \notin U$, we have $\pi(x) \neq 0$. By Corollary E.8, there exists $\varphi \in(X / U)^{\prime}$ such that $\varphi(\pi(x)) \neq 0$. Then $x^{\prime}:=\varphi \circ \pi \in X^{\prime}$ is a functional we are looking for.

Definition E.12. A linear subspace $U$ of a normed space $X$ is called complemented if there exists a projection $P \in \mathscr{L}(X)$ such that $\operatorname{ran} P=U$.

Remark E.13. If $P$ is a projection (that means, if $P^{2}=P$ ), then $Q=I-P$ is also a projection and $\operatorname{ran} P=\operatorname{ker} Q$. Hence, if $P$ is a bounded projection, then $\operatorname{ran} P$ is
necessarily closed. Thus, a necessary condition for $U$ to be complemented is that $U$ is closed.

Corollary E.14. Every finite dimensional subspace of a normed space is complemented.

Proof. Let $U$ be a finite dimensional subspace of a normed space $X$. Let $\left(b_{i}\right)_{1 \leq i \leq N}$ be a basis of $U$. By Corollary E.11, there exist functionals $x_{i}^{\prime} \in X^{\prime}$ such that

$$
\left\langle x_{i}^{\prime}, b_{j}\right\rangle=\left\{\begin{array}{l}
1 \text { if } i=j \\
0 \text { otherwise }
\end{array}\right.
$$

Let $P: X \rightarrow X$ be defined by

$$
P x:=\sum_{i=1}^{N}\left\langle x_{i}^{\prime}, x\right\rangle b_{i}, \quad x \in X .
$$

Then $P b_{i}=b_{i}$ for every $1 \leq i \leq N$, and thus $P^{2}=P$, that is, $P$ is a projection. Moreover, $\operatorname{ran} P=U$ by construction. By the estimate

$$
\begin{aligned}
\|P x\| & \leq \sum_{i=1}^{N}\left|\left\langle x_{i}^{\prime}, x\right\rangle\right|\left\|b_{i}\right\| \\
& \leq\left(\sum_{i=1}^{N}\left\|x_{i}^{\prime}\right\|\left\|b_{i}\right\|\right)\|x\|,
\end{aligned}
$$

the projection $P$ is bounded.
The following lemma which does not depend on the Hahn-Banach theorem is stated for completeness.

Lemma E.15. In a Hilbert space every closed linear subspace is complemented.
Proof. Take the orthogonal projection onto the closed subspace as a possible projection.

Corollary E.16. If $X$ is a normed space such that $X^{\prime}$ is separable, then $X$ is separable, too.

Proof. Let $D^{\prime}=\left\{x_{n}^{\prime}: n \in \mathbb{N}\right\}$ be a dense subset of the unit sphere of $X^{\prime}$. For every $n \in \mathbb{N}$ we choose an element $x_{n} \in X$ such that $\left\|x_{n}\right\| \leq 1$ and $\left|\left\langle x_{n}^{\prime}, x_{n}\right\rangle\right| \geq \frac{1}{2}$. We claim that $D:=\operatorname{span}\left\{x_{n}: n \in \mathbb{N}\right\}$ is dense in $X$. If this was not true, i.e. if $\bar{D} \neq X$, then, by Corollary E.11, we find an element $x^{\prime} \in X^{\prime} \backslash\{0\}$ such that $x^{\prime}\left(x_{n}\right)=0$ for every $n \in \mathbb{N}$. We may without loss of generality assume that $\left\|x^{\prime}\right\|=1$. Since $D^{\prime}$ is dense in the unit sphere of $X^{\prime}$, we find $n_{0} \in \mathbb{N}$ such that $\left\|x^{\prime}-x_{n_{0}}^{\prime}\right\| \leq \frac{1}{4}$. But then

$$
\frac{1}{2} \leq\left|\left\langle x_{n_{0}}^{\prime}, x_{n_{0}}\right\rangle\right|=\left|\left\langle x_{n_{0}}^{\prime}-x^{\prime}, x_{n_{0}}\right\rangle\right| \leq\left\|x_{n_{0}}^{\prime}-x^{\prime}\right\|\left\|x_{n_{0}}\right\| \leq \frac{1}{4}
$$

which is a contradiction. Hence, $\bar{D}=X$ and $X$ is separable by Lemma D. 21 of Chapter D.

## E. 2 Weak* convergence and the theorem of Banach-Alaoglu

Definition E.17. Let $X$ be a Banach space. We say that a sequence $\left(x_{n}^{\prime}\right) \subset X^{\prime}$ converges weak ${ }^{*}$ to some element $x^{\prime} \in X^{\prime}$ if for every $x \in X$ one has $\lim _{n \rightarrow \infty}\left\langle x_{n}^{\prime}, x\right\rangle=$ $\left\langle x^{\prime}, x\right\rangle$. We write $x_{n}^{\prime} \xrightarrow{\text { weak* }} x^{\prime}$ if $\left(x_{n}^{\prime}\right)$ converges weak ${ }^{*}$ to $x^{\prime}$.
Theorem E. 18 (Banach-Alaoglu). Let X be a separable Banach space. Then every bounded sequence $\left(x_{n}^{\prime}\right) \subset X^{\prime}$ admits a weak ${ }^{*}$ convergent subsequence, that is, there exists $x^{\prime} \in X^{\prime}$ and there exists a subsequence $\left(x_{n_{k}}^{\prime}\right)$ of $\left(x_{n}^{\prime}\right)$ such that $x_{n_{k}}^{\prime} \xrightarrow{\text { weak* }} x^{\prime}$.
Proof. As in the proof of the Arzela-Ascoli theorem (Theorem B.43) and the theorem about weak sequential compactness of the unit ball in Hilbert spaces (Theorem D.57), we use Cantor's diagonal sequence argument. Let $\left(x_{n}^{\prime}\right)$ be a bounded sequence in $X^{\prime}$.

Since $X$ is separable by assumption, we can choose a dense sequence $\left(x_{m}\right) \subset X$. Since $\left(\left\langle x_{n}^{\prime}, x_{1}\right\rangle\right)$ is bounded by the boundedness of $\left(x_{n}^{\prime}\right)$, there exists a subsequence $\left(x_{\varphi_{1}(n)}^{\prime}\right)$ of $\left(x_{n}^{\prime}\right)\left(\varphi_{1}: \mathbb{N} \rightarrow \mathbb{N}\right.$ is increasing, unbounded) such that

$$
\lim _{n \rightarrow \infty}\left\langle x_{\varphi_{1}(n)}^{\prime}, x_{1}\right\rangle \text { exists. }
$$

Similarly, there exists a subsequence $\left(x_{\varphi_{2}(n)}^{\prime}\right)$ of $\left(x_{\varphi_{1}(n)}^{\prime}\right)$ such that

$$
\lim _{n \rightarrow \infty}\left\langle x_{\varphi_{2}(n)}^{\prime}, x_{2}\right\rangle \text { exists. }
$$

Note that for this subsequence, we also have that

$$
\lim _{n \rightarrow \infty}\left\langle x_{\varphi_{2}(n)}^{\prime}, x_{1}\right\rangle \text { exists. }
$$

Iterating this argument, we find a subsequence $\left(x_{\varphi_{3}(n)}^{\prime}\right)$ of $\left(x_{\varphi_{2}(n)}^{\prime}\right)$ and finally for every $m \in \mathbb{N}, m \geq 2$, a subsequence $\left(x_{\varphi_{m}(n)}^{\prime}\right)$ of $\left(x_{\varphi_{m-1}(n)}^{\prime}\right)$ such that

$$
\lim _{n \rightarrow \infty}\left\langle x_{\varphi_{m}(n)}^{\prime}, x_{j}\right\rangle \text { exists for every } 1 \leq j \leq m
$$

Let $\left(y_{n}^{\prime}\right):=\left(x_{\varphi_{n}(n)}^{\prime}\right)$ be the 'diagonal sequence'. Then $\left(y_{n}^{\prime}\right)$ is a subsequence of $\left(x_{n}^{\prime}\right)$ and

$$
\lim _{n \rightarrow \infty}\left\langle y_{n}^{\prime}, x_{m}\right\rangle \text { exists for every } m \in \mathbb{N}
$$

By Lemma D. 58 of Chapter D, there exists $x^{\prime} \in X^{\prime}$ such that

$$
\lim _{n \rightarrow \infty}\left\langle y_{n}^{\prime}, x\right\rangle=\left\langle x^{\prime}, x\right\rangle \text { for every } x \in X
$$

This is the claim.

## E. 3 Weak convergence and reflexivity

Given a normed space $X$, we call $X^{\prime \prime}:=\left(X^{\prime}\right)^{\prime}=\mathscr{L}\left(X^{\prime}, \mathbb{K}\right)$ the bidual of $X$.
Lemma E.19. Let $X$ be a normed space. Then the mapping

$$
\begin{aligned}
J: X & \rightarrow X^{\prime \prime}, \\
x & \mapsto\left(x^{\prime} \mapsto\left\langle x^{\prime}, x\right\rangle\right),
\end{aligned}
$$

is well defined and isometric.
Proof. The linearity of $x^{\prime} \mapsto\left\langle x^{\prime}, x\right\rangle$ is clear, and from the inequality

$$
\left|J x\left(x^{\prime}\right)\right|=\left|\left\langle x^{\prime}, x\right\rangle\right| \leq\left\|x^{\prime}\right\|\|x\|
$$

follows that $J x \in X^{\prime \prime}$ (i.e. $J$ is well defined) and $\|J x\| \leq\|x\|$. The fact that $J$ is isometric follows from Corollary E.8.

Definition E.20. A normed space $X$ is called reflexive if the isometry $J$ from Lemma E. 19 is surjective, i.e. if $J X=X^{\prime \prime}$. In other words: a normed space $X$ is reflexive if for every $x^{\prime \prime} \in X^{\prime \prime}$ there exists $x \in X$ such that

$$
\left\langle x^{\prime \prime}, x^{\prime}\right\rangle=\left\langle x^{\prime}, x\right\rangle \text { for all } x^{\prime} \in X^{\prime} .
$$

Remark E.21. If a normed space is reflexive then $X$ and $X^{\prime \prime}$ are isometrically isomorphic (via the operator $J$ ). Since $X^{\prime \prime}$ is always complete, a reflexive space is necessarily a Banach space.

Note that it can happen that $X$ and $X^{\prime \prime}$ are isomorphic without $X$ being reflexive (the example of such a Banach space is however quite involved). We point out that reflexivity means that the special operator $J$ is an isomorphism.

Lemma E.22. Every Hilbert space is reflexive.
Proof. By the Theorem of Riesz-Fréchet, we may identify $H$ with its dual $H^{\prime}$ and thus also $H$ with its bidual $H^{\prime \prime}$. The identification is done via the scalar product. It is an exercise to show that this identification of $H$ with $H^{\prime \prime}$ coincides with the mapping $J$ from Lemma E. 19 .

Remark E.23. It should be noted that for complex Hilbert spaces, the identification of $H$ with its dual $H^{\prime}$ is only antilinear, but after the second identification ( $H^{\prime}$ with $H^{\prime \prime}$ ) it turns out that the identification of $H$ with $H^{\prime \prime}$ is linear.

Lemma E.24. Every finite dimensional Banach space is reflexive.

Proof. It suffices to remark that if $X$ is finite dimensional, then

$$
\operatorname{dim} X=\operatorname{dim} X^{\prime}=\operatorname{dim} X^{\prime \prime}<\infty .
$$

Surjectivity of the mapping $J$ (which is always injective) thus follows from linear algebra.

Theorem E.25. The space $L^{p}(\Omega)$ is reflexive if $1<p<\infty((\Omega, \mathscr{A}, \mu)$ being an arbitrary measure space).

We will actually only prove the following special case.
Theorem E.26. The spaces $l^{p}$ are reflexive if $1<p<\infty$.
The proof of Theorem E. 26 is based on the following lemma.
Lemma E.27. Let $1 \leq p<\infty$ and let $q:=\frac{p}{p-1}$ be the conjugate exponent so that $\frac{1}{p}+\frac{1}{q}=1$. Then the operator

$$
\begin{aligned}
T: l^{q} & \rightarrow\left(l^{p}\right)^{\prime} \\
\left(a_{n}\right) & \mapsto\left(\left(x_{n}\right) \mapsto \sum_{n} a_{n} x_{n}\right),
\end{aligned}
$$

is an isometric isomorphism, i.e. $\left(l^{p}\right)^{\prime}=l^{q}$.
Proof. Linearity of $T$ is obvious. Assume first $p>1$, so that $q<\infty$. Note that for every $a:=\left(a_{n}\right) \in l^{q} \backslash\{0\}$ the sequence $\left(x_{n}\right):=\left(c \overline{a_{n}}\left|a_{n}\right|^{q-2}\right)\left(c=\|a\|_{q}^{-q / p}\right)$ belongs to $l^{p}$ and

$$
\|x\|_{p}^{p}=\|a\|_{q}^{-q} \sum_{n}\left|a_{n}\right|^{(q-1) p}=1
$$

This particular $x \in l^{p}$ shows that

$$
\|T a\|_{(l p)^{\prime}} \geq \sum_{n} a_{n} x_{n}=\|a\|_{q}^{-q / p} \sum_{n}\left|a_{n}\right|^{q}=\|a\|_{q}^{q(p-1) / p}=\|a\|_{q} .
$$

On the other hand, by Hölder's inequality,

$$
\|T a\|_{(l)^{\prime}}=\sup _{\|x\|_{p} \leq 1}\left|\sum_{n} a_{n} x_{n}\right| \leq\|a\|_{q}
$$

so that $T$ is isometric in the case $p \in(1, \infty)$. The case $p=1$ is very similar and will be omitted.

In order to show that $T$ is surjective, let $\varphi \in\left(l^{p}\right)^{\prime}$. Denote by $e_{n}$ the $n$-th unit vector in $l^{p}$, and let $a_{n}:=\varphi\left(e_{n}\right)$. If $p=1$, then $\left(a_{n}\right) \in l^{\infty}=l^{q}$ by the trivial estimate

$$
\left|a_{n}\right|=\left|\varphi\left(e_{n}\right)\right| \leq\|\varphi\|\left\|e_{n}\right\|_{1}=\|\varphi\|
$$

If $p>1$, then we may argue as follows. For every $N \in \mathbb{N}$,

$$
\begin{aligned}
\sum_{n=1}^{N}\left|a_{n}\right|^{q} & =\sum_{n=1}^{N} a_{n} \overline{a_{n}}\left|a_{n}\right|^{q-2} \\
& =\varphi\left(\sum_{n=1}^{N} \overline{a_{n}}\left|a_{n}\right|^{q-2} e_{n}\right) \\
& \leq\|\varphi\|\left(\sum_{n=1}^{N}\left|a_{n}\right|^{(q-1) p}\right)^{\frac{1}{p}} \\
& =\|\varphi\|\left(\sum_{n=1}^{N}\left|a_{n}\right|^{q}\right)^{\frac{1}{p}}
\end{aligned}
$$

which is equivalent to

$$
\left(\sum_{n=1}^{N}\left|a_{n}\right|^{q}\right)^{1-\frac{1}{p}}=\left(\sum_{n=1}^{N}\left|a_{n}\right|^{q}\right)^{\frac{1}{q}} \leq\|\varphi\| .
$$

Since the right-hand side of this inequality does not depend on $N \in \mathbb{N}$, we obtain that $a:=\left(a_{n}\right) \in l^{q}$ and $\|a\|_{q} \leq\|\varphi\|$.

Next, observe that for every $x \in l^{p}$ one has

$$
x=\sum_{n} x_{n} e_{n}=\lim _{N \rightarrow \infty} \sum_{n=1}^{N} x_{n} e_{n}
$$

the series converging in $l^{p}$ (here we need the restriction $p<\infty!$ ). Hence, for every $x \in l^{p}$, by the boundedness of $\varphi$,

$$
\begin{aligned}
\varphi(x) & =\lim _{N \rightarrow \infty} \varphi\left(\sum_{n=1}^{N} x_{n} e_{n}\right) \\
& =\lim _{N \rightarrow \infty} \sum_{n=1}^{N} x_{n} a_{n} \\
& =\sum_{n} x_{n} a_{n} \\
& =T a(x)
\end{aligned}
$$

Hence, $T$ is surjective.
Proof (Proof of Theorem E.26). By Lemma E.27, we may identify $\left(l^{p}\right)^{\prime}$ with $l^{q}$ and, if $1<p<\infty(!)$, also $\left(l^{p}\right)^{\prime \prime}=\left(l^{q}\right)^{\prime}$ with $l^{p}$. One just has to notice that this identification of $l^{p}$ with $\left(l^{p}\right)^{\prime \prime}=l^{p}$ (the identity map on $l^{p}$ ) coincides with the operator $J$ from Lemma E.19, so that $l^{p}$ is reflexive if $1<p<\infty$.
Lemma E.28. The spaces $l^{1}, L^{1}(\Omega)\left(\Omega \subset \mathbb{R}^{N}\right)$ and $C([0,1])$ are not reflexive.
Proof. For every $t \in[0,1]$, let $\delta_{t} \in C([0,1])^{\prime}$ be defined by

$$
\left\langle\delta_{t}, f\right\rangle:=f(t), \quad f \in C([0,1])
$$

Then $\left\|\delta_{t}\right\|=1$ and whenever $t \neq s$, then

$$
\left\|\delta_{t}-\delta_{s}\right\|=2
$$

In particular, the uncountably many balls $B\left(\delta_{t}, \frac{1}{2}\right)(t \in[0,1])$ are mutually disjoint so that $C([0,1])^{\prime}$ is not separable.

Now, if $C([0,1])$ were reflexive, then $C([0,1])^{\prime \prime}=C([0,1])$ would be separable (since $C([0,1])$ is separable), and then, by Corollary E.16, $C([0,1])^{\prime}$ would be separable; a contradiction to what has been said before. This proves that $C([0,1])$ is not reflexive.

The cases of $l^{1}$ and $L^{1}(\Omega)$ are proved similarly. They are separable Banach spaces with nonseparable dual.

Theorem E.29. Every closed subspace of a reflexive Banach space is reflexive.
Proof. Let $X$ be a reflexive Banach space, and let $U \subset X$ be a closed subspace. Let $u^{\prime \prime} \in U^{\prime \prime}$. Then the mapping $x^{\prime \prime}: X^{\prime} \rightarrow \mathbb{K}$ defined by

$$
\left\langle x^{\prime \prime}, x^{\prime}\right\rangle=\left\langle u^{\prime \prime},\left.x^{\prime}\right|_{U}\right\rangle, \quad x^{\prime} \in X^{\prime}
$$

is linear and bounded, i.e. $x^{\prime \prime} \in X^{\prime \prime}$. By reflexivity of $X$, there exists $x \in X$ such that

$$
\begin{equation*}
\left\langle x^{\prime}, x\right\rangle=\left\langle u^{\prime \prime},\left.x^{\prime}\right|_{U}\right\rangle, \quad x^{\prime} \in X^{\prime} \tag{E.5}
\end{equation*}
$$

Assume that $x \notin U$. Then, by Corollary E.9, there exists $x^{\prime} \in X^{\prime}$ such that $\left.x^{\prime}\right|_{U}=0$ and $\left\langle x^{\prime}, x\right\rangle \neq 0$; a contradiction to the last equality. Hence, $x \in U$. We need to show that

$$
\begin{equation*}
\left\langle u^{\prime \prime}, u^{\prime}\right\rangle=\left\langle u^{\prime}, x\right\rangle, \forall u^{\prime} \in U^{\prime} \tag{E.6}
\end{equation*}
$$

However, if $u^{\prime} \in U^{\prime}$, then, by Hahn-Banach we can choose an extension $x^{\prime} \in X^{\prime}$, i.e. $\left.x^{\prime}\right|_{U}=u^{\prime}$. The equation (E.6) thus follows from (E.5).

Corollary E.30. The Sobolev spaces $W^{k, p}(\Omega)\left(\Omega \subset \mathbb{R}^{N}\right.$ open) are reflexive if $1<$ $p<\infty, k \in \mathbb{N}$.

Proof. For example, for $k=1$, the operator

$$
\begin{aligned}
T: W^{1, p}(\Omega) & \rightarrow L^{p}(\Omega)^{1+N} \\
u & \mapsto\left(u, \frac{\partial u}{\partial x_{1}}, \ldots, \frac{\partial u}{\partial x_{N}}\right)
\end{aligned}
$$

is isometric, so that we may consider $W^{1, p}(\Omega)$ as a closed subspace of $L^{p}(\Omega)^{1+N}$ which is reflexive by Theorem E.25. The claim follows from Theorem E.29.

Corollary E.31. A Banach space is reflexive if and only if its dual is reflexive.
Proof. Assume that the Banach space $X$ is reflexive. Let $x^{\prime \prime \prime} \in X^{\prime \prime \prime}$ (the tridual!). Then the mapping $x^{\prime}: X \rightarrow \mathbb{K}$ defined by

$$
\left\langle x^{\prime}, x\right\rangle:=\left\langle x^{\prime \prime \prime}, J_{X}(x)\right\rangle, \quad x \in X,
$$

is linear and bounded, i.e. $x^{\prime} \in X^{\prime}$ (here $J_{X}$ denotes the isometry $X \rightarrow X^{\prime \prime}$ ). Let $x^{\prime \prime} \in$ $X^{\prime \prime}$ be arbitrary. Since $X$ is reflexive, there exists $x \in X$ such that $J_{X} x=x^{\prime \prime}$. Hence,

$$
\left\langle x^{\prime \prime \prime}, x^{\prime \prime}\right\rangle=\left\langle x^{\prime \prime \prime}, J_{X} x\right\rangle=\left\langle x^{\prime}, x\right\rangle=\left\langle x^{\prime \prime}, x^{\prime}\right\rangle
$$

which proves that $J_{X^{\prime}} x^{\prime}=x^{\prime \prime \prime}$, i.e. the isometry $J_{X^{\prime}}: X^{\prime} \rightarrow X^{\prime \prime \prime}$ is surjective. Hence, $X^{\prime}$ is reflexive.

On the other hand, assume that $X^{\prime}$ is reflexive. Then $X^{\prime \prime}$ is reflexive by the preceeding argument, and therefore $X$ (considered as a closed subspace of $X^{\prime \prime}$ via the isometry $J$ ) is reflexive by Theorem E.29.

Definition E.32. Let $X$ be a normed space. We say that a sequence $\left(x_{n}\right) \subset X$ converges weakly to some $x \in X$ if

$$
\lim _{n \rightarrow \infty}\left\langle x^{\prime}, x_{n}\right\rangle=\left\langle x^{\prime}, x\right\rangle \text { for every } x^{\prime} \in X^{\prime}
$$

Notations: if $\left(x_{n}\right)$ converges weakly to $x$, then we write $x_{n} \rightharpoonup x, w-\lim _{n \rightarrow \infty} x_{n}=x$, $x_{n} \rightarrow x$ in $\sigma\left(X, X^{\prime}\right)$, or $x_{n} \rightarrow x$ weakly.

Theorem E.33. In a reflexive Banach space every bounded sequence admits a weakly convergent subsequence.

Proof. Let $\left(x_{n}\right)$ be a bounded sequence in a reflexive Banach space $X$. We first assume that $X$ is separable. Then $X^{\prime \prime}$ is separable by reflexivity, and $X^{\prime}$ is separable by Corollary E.16. Let $\left(x_{m}^{\prime}\right) \subset X^{\prime}$ be a dense sequence.

Since $\left(\left\langle x_{1}^{\prime}, x_{n}\right\rangle\right)$ is bounded by the boundedness of $\left(x_{n}\right)$, there exists a subsequence $\left(x_{\varphi_{1}(n)}\right)$ of $\left(x_{n}\right)\left(\varphi_{1}: \mathbb{N} \rightarrow \mathbb{N}\right.$ is increasing, unbounded) such that

$$
\lim _{n \rightarrow \infty}\left\langle x_{1}^{\prime}, x_{\varphi_{1}(n)}\right\rangle \text { exists. }
$$

Similarly, there exists a subsequence $\left(x_{\varphi_{2}(n)}\right)$ of $\left(x_{\varphi_{1}(n)}\right)$ such that

$$
\lim _{n \rightarrow \infty}\left\langle x_{2}^{\prime}, x_{\varphi_{2}(n)}\right\rangle \text { exists. }
$$

Note that for this subsequence, we also have that

$$
\lim _{n \rightarrow \infty}\left\langle x_{1}^{\prime}, x_{\varphi_{2}(n)}\right\rangle \text { exists. }
$$

Iterating this argument, we find a subsequence $\left(x_{\varphi_{3}(n)}\right)$ of $\left(x_{\varphi_{2}(n)}\right)$ and finally for every $m \in \mathbb{N}, m \geq 2$, a subsequence $\left(x_{\varphi_{m}(n)}\right)$ of $\left(x_{\varphi_{m-1}(n)}\right)$ such that

$$
\lim _{n \rightarrow \infty}\left\langle x_{j}^{\prime}, x_{\varphi_{m}(n)}\right\rangle \text { exists for every } 1 \leq j \leq m
$$

Let $\left(y_{n}\right):=\left(x_{\varphi_{n}(n)}\right)$ be the 'diagonal sequence'. Then $\left(y_{n}\right)$ is a subsequence of $\left(x_{n}\right)$ and

$$
\lim _{n \rightarrow \infty}\left\langle x_{m}^{\prime}, y_{n}\right\rangle \text { exists for every } m \in \mathbb{N}
$$

By Lemma D. 58 of Chapter D, there exists $x^{\prime \prime} \in X^{\prime \prime}$ such that

$$
\lim _{n \rightarrow \infty}\left\langle x^{\prime}, y_{n}\right\rangle=\left\langle x^{\prime}, x^{\prime \prime}\right\rangle \text { for every } x^{\prime} \in X^{\prime}
$$

Since $X$ is reflexive, there exists $x \in X$ such that $J x=x^{\prime \prime}$. For this $x$, we have by definition of $J$

$$
\lim _{n \rightarrow \infty}\left\langle x^{\prime}, y_{n}\right\rangle=\left\langle x^{\prime}, x\right\rangle \text { exists for every } x^{\prime} \in X^{\prime}
$$

i.e. $\left(y_{n}\right)$ converges weakly to $x$.

If $X$ is not separable as we first assumed, then one may replace $X$ by $\tilde{X}:=$ $\overline{\operatorname{span}}\left\{x_{n}: n \in \mathbb{N}\right\}$ which is separable. By the above, there exists $x \in \tilde{X}$ and a subsequence of $\left(x_{n}\right)$ (which we denote again by $\left(x_{n}\right)$ ) such that for every $\tilde{x}^{\prime} \in \tilde{X}^{\prime}$,

$$
\lim _{n \rightarrow \infty}\left\langle\tilde{x}^{\prime}, x_{n}\right\rangle=\left\langle\tilde{x}^{\prime}, x\right\rangle
$$

i.e. $\left(x_{n}\right)$ converges weakly in $\tilde{X}$. If $x^{\prime} \in X^{\prime}$, then $\left.x^{\prime}\right|_{\tilde{X}} \in \tilde{X}^{\prime}$, and it follows easily that the sequence $\left(x_{n}\right)$ also converges weakly in $X$ to the element $x$.

## E. 4 * Minimization of convex functionals

Theorem E. 34 (Hahn-Banach; separation of convex sets). Let $X$ be a Banach space, $K \subset X$ a closed, nonempty, convex subset, and $x_{0} \in X \backslash K$. Then there exists $x^{\prime} \in X^{\prime}$ and $\varepsilon>0$ such that

$$
\operatorname{Re}\left\langle x^{\prime}, x\right\rangle+\varepsilon \leq \operatorname{Re}\left\langle x^{\prime}, x_{0}\right\rangle, \quad x \in K
$$

Lemma E.35. Let $K$ be an open, nonempty, convex subset of a Banach space $X$ such that $0 \in K$. Define the Minkowski functional $p: X \rightarrow \mathbb{R}$ by

$$
p(x)=\inf \left\{\lambda>0: \frac{x}{\lambda} \in K\right\} .
$$

Then $p$ is sublinear, there exists $M \geq 0$ such that

$$
p(x) \leq M\|x\|, \quad x \in X
$$

and $K=\{x \in X: p(x)<1\}$.
Proof. Since $B(0, r) \subset K$ for some $r>0$, we find that

$$
p(x) \leq \frac{1}{r}\|x\| \text { for every } x \in X
$$

The property $p(\alpha x)=\alpha p(x)$ for every $\alpha>0$ and every $x \in X$ is obvious.

Next, if $p(x)<1$, then there exists $\lambda \in(0,1)$ such that $x / \lambda \in K$. Hence, by convexity, $x=\lambda \frac{x}{\lambda}=\lambda \frac{x}{\lambda}+(1-\lambda) 0 \in K$. On the other hand, if $x \in K$, then $(1+\varepsilon) x \in$ $K$, since $K$ is open. Hence, $p(x) \leq(1+\varepsilon)^{-1}<1$, so that $K=\{x \in X: p(x)<1\}$.

Let finally $x, y \in X$. Then for every $\varepsilon>0, x /(p(x)+\varepsilon) \in K$ and $y /(p(y)+\varepsilon) \in K$. In particular, for every $t \in[0,1]$,

$$
\frac{t}{p(x)+\varepsilon} x+\frac{1-t}{p(y)+\varepsilon} y \in K .
$$

Setting $t=(p(x)+\varepsilon) /(p(x)+p(y)+2 \varepsilon)$, one finds that

$$
\frac{x+y}{p(x)+p(y)+2 \varepsilon} \in K
$$

so that $p(x+y) \leq p(x)+p(y)+2 \varepsilon$. Since $\varepsilon>0$ was arbitrary, we find $p(x+y) \leq$ $p(x)+p(y)$. The claim is proved.

Proof (Proof of Theorem E.34). We prove the theorem for the case when $X$ is a real Banach space. The complex case is proved similarly.

We may without loss of generality assume that $0 \in K$; it suffices to translate $K$ and $x_{0}$ for this. Since $x_{0} \notin K$ and since $K$ is closed, we find that $d:=\operatorname{dist}\left(x_{0}, K\right)>0$. Put

$$
K_{d}:=\{x \in X: \operatorname{dist}(x, K)<d / 2\}
$$

so that $K_{d}$ is an open, convex subset such that $0 \in K_{d}$. Let $p$ be the corresponding Minkowski functional (see Lemma E.35).

Define on the one-dimensional subspace $U:=\left\{\lambda x_{0}: \lambda \in \mathbb{R}\right\}$ the functional $u^{\prime}$ : $U \rightarrow \mathbb{R}$ by $\left\langle u^{\prime}, \lambda x_{0}\right\rangle=\lambda$. Then

$$
\left\langle u^{\prime}, u\right\rangle \leq p(u), \quad u \in U .
$$

By the Hahn-Banach theorem E.3, there exists a linear extension $x^{\prime}: X \rightarrow \mathbb{R}$ such that

$$
\begin{equation*}
\left\langle x^{\prime}, x\right\rangle \leq p(x), \quad x \in X \tag{E.7}
\end{equation*}
$$

In particular, by Lemma E.35,

$$
\left|\left\langle x^{\prime}, x\right\rangle\right| \leq M\|x\|
$$

so that $x^{\prime} \in X^{\prime}$ and $\left\|x^{\prime}\right\| \leq M$. By construction, $\left\langle x^{\prime}, x_{0}\right\rangle=1$. Moreover, by (E.7) and Lemma E.35, $\left\langle x^{\prime}, x\right\rangle<1$ for every $x \in K \subset K_{d}$, so that

$$
\left\langle x^{\prime}, x\right\rangle \leq\left\langle x^{\prime}, x_{0}\right\rangle(=1), \quad x \in K_{d}
$$

Replacing the above argument with $\left(1-\varepsilon^{\prime}\right) x_{0}$ instead of $x_{0}$ (where $\varepsilon^{\prime}>0$ is chosen so small that $\left.\left(1-\varepsilon^{\prime}\right) x_{0} \notin K_{d}\right)$, we find that

$$
\left\langle x^{\prime}, x\right\rangle+\varepsilon^{\prime}\left\langle x^{\prime}, x_{0}\right\rangle \leq\left\langle x^{\prime}, x_{0}\right\rangle, \quad x \in K \subset K_{d}
$$

and putting $\varepsilon:=\varepsilon^{\prime}=\varepsilon^{\prime}\left\langle x^{\prime}, x_{0}\right\rangle>0$ yields the claim.
Corollary E.36. Let $X$ be a Banach space and $K \subset X$ a closed, convex subset (closed for the norm topology). If $\left(x_{n}\right) \subset K$ converges weakly to some $x \in X$, then $x \in K$.

Proof. Assume the contrary, i.e. $x \notin K$. By the Hahn-Banach theorem (Theorem E.34), there exist $x^{\prime} \in X^{\prime}$ and $\varepsilon>0$ such that

$$
\operatorname{Re}\left\langle x^{\prime}, x_{n}\right\rangle+\varepsilon \leq \operatorname{Re}\left\langle x^{\prime}, x\right\rangle \text { for every } n \in \mathbb{N}
$$

a contradiction to the assumption that $x_{n} \rightharpoonup x$.
A function $f: K \rightarrow \mathbb{R}$ on a convex subset $K$ of a Banach space $X$ is called convex if for every $x, y \in K$, and every $t \in[0,1]$,

$$
\begin{equation*}
f(t x+(1-t) y) \leq t f(x)+(1-t) f(y) \tag{E.8}
\end{equation*}
$$

Corollary E.37. Let $X$ be a Banach space, $K \subset X$ a closed, convex subset, and $f: K \rightarrow \mathbb{R}$ a continuous, convex function. If $\left(x_{n}\right) \subset K$ converges weakly to $x \in K$, then

$$
f(x) \leq \liminf _{n \rightarrow \infty} f\left(x_{n}\right)
$$

Proof. For every $l \in \mathbb{R}$, the set $K_{l}:=\{x \in K: f(x) \leq l\}$ is closed (by continuity of $f$ ) and convex (by convexity of $f$ ). After extracting a subsequence, if necessary, we may assume that $l:=\liminf _{n \rightarrow \infty} f\left(x_{n}\right)=\lim _{n \rightarrow \infty} f\left(x_{n}\right)$. Then for every $\varepsilon>0$ the sequence $\left(x_{n}\right)$ is eventually in $K_{l+\varepsilon}$, i.e. except for finitely many $x_{n}$, the sequence $\left(x_{n}\right)$ lies in $K_{l+\varepsilon}$. Hence, by Corollary E. $36, x \in K_{l+\varepsilon}$, which means that $f(x) \leq l+\varepsilon$. Since $\varepsilon>0$ was arbitrary, the claim follows.

Theorem E.38. Let $X$ be a reflexive Banach space, $K \subset X$ a closed, convex, nonempty subset, and $f: K \rightarrow \mathbb{R}$ a continuous, convex function such that

$$
\lim _{\substack{\|x\| \rightarrow \infty \\ x \in K}} f(x)=+\infty(\text { coercivity })
$$

Then there exists $x_{0} \in K$ such that

$$
f\left(x_{0}\right)=\inf \{f(x): x \in K\}>-\infty .
$$

Proof. Let $\left(x_{n}\right) \subset K$ be such that $\lim _{n \rightarrow \infty} f\left(x_{n}\right)=\inf \{f(x): x \in K\}$. By the coercivity assumption on $f$, the sequence $\left(x_{n}\right)$ is bounded. Since $X$ is reflexive, there exists a weakly convergent subsequence (Theorem E.33); we denote by $x_{0}$ the limit. By Corollary E. $36, x_{0} \in K$. By Corollary E.37,

$$
f\left(x_{0}\right) \leq \lim _{n \rightarrow \infty} f\left(x_{n}\right)=\inf \{f(x): x \in K\}
$$

The claim is proved.

Remark E.39. Theorem E. 38 remains true if $f$ is only lower semicontinuous, i.e. if

$$
\liminf _{n \rightarrow \infty} f\left(x_{n}\right) \geq f(x)
$$

for every convergent $\left(x_{n}\right) \subset K$ with $x=\lim x_{n}$. In fact, already Corollary E. 37 remains true if $f$ is only lower semicontinuous (and then Corollary E. 37 says that lower semicontinuity of a convex function in the norm topology implies lower semicontinuity in the weak topology). It suffices for example to remark that the sets $K_{l}:=\{f \leq l\}(l \in \mathbb{R})$ are closed as soon as $f$ is lower semicontinuous.

## E. 5 * The von Neumann minimax theorem

In the following theorem, we call a function $f: K \rightarrow \mathbb{R}$ on a convex subset $K$ of a Banach space $X$ concave if $-f$ is convex, or, equivalently, if for every $x, y \in K$ and every $t \in[0,1]$,

$$
\begin{equation*}
f(t x+(1-t) y) \geq t f(x)+(1-t) f(y) \tag{E.9}
\end{equation*}
$$

A function $f: K \rightarrow \mathbb{R}$ is called strictly convex (resp. strictly concave) if for every $x$, $y \in K, x \neq y, f(x)=f(y)$ the inequality in (E.8) (resp. (E.9)) is strict for $t \in(0,1)$.

Theorem E. 40 (von Neumann). Let $K$ and L be two closed, bounded, nonempty, convex subsets of reflexive Banach spaces $X$ and $Y$, respectively. Let $f: K \times L \rightarrow \mathbb{R}$ be a continuous function such that

$$
\begin{aligned}
& x \mapsto f(x, y) \text { is strictly convex for every } y \in L, \text { and } \\
& y \mapsto f(x, y) \text { is concave for every } x \in K .
\end{aligned}
$$

Then there exists $(\bar{x}, \bar{y}) \in K \times L$ such that

$$
\begin{equation*}
f(\bar{x}, y) \leq f(\bar{x}, \bar{y}) \leq f(x, \bar{y}) \text { for every } x \in K, y \in L \tag{E.10}
\end{equation*}
$$

Remark E.41. A point $(\bar{x}, \bar{y}) \in K \times L$ satisfying (E.10) is called a saddle point of $f$.
A saddle point is a point of equilibrium in a two-person zero-sum game in the following sense: If the player controlling the strategy $x$ modifies his strategy when the second player plays $\bar{y}$, he increases his loss; hence, it is his interest to play $\bar{x}$. Similarly, if the player controlling the strategy $y$ modifies his strategy when the first player plays $\bar{x}$, he diminishes his gain; thus it is in his interest to play $\bar{y}$. This property of equilibrium of saddle points justifies their use as a (reasonable) solution in a two-person zero-sum game ([Aubin (1979)]).

Proof. Define the function $F: L \rightarrow \mathbb{R}$ by $F(y):=\inf _{x \in K} f(x, y)(y \in L)$. By Theorem E.38, for every $y \in L$ there exists $x \in K$ such that $F(y)=f(x, y)$. By strict convexity, this element $x$ is uniquely determined. We denote $x:=\Phi(y)$ and thus obtain

$$
\begin{equation*}
F(y)=\inf _{x \in K} f(x, y)=f(\Phi(y), y), \quad y \in L \tag{E.11}
\end{equation*}
$$

By concavity of the function $y \mapsto f(x, y)$ and by the definition of $F$, for every $y_{1}$, $y_{2} \in L$ and every $t \in[0,1]$,

$$
\begin{aligned}
F\left(t y_{1}+(1-t) y_{2}\right) & =f\left(\Phi\left(t y_{1}+(1-t) y_{2}\right), t y_{1}+(1-t) y_{2}\right) \\
& \geq t f\left(\Phi\left(t y_{1}+(1-t) y_{2}\right), y_{1}\right)+(1-t) f\left(\Phi\left(t y_{1}+(1-t) y_{2}\right), y_{2}\right) \\
& \geq t F\left(y_{1}\right)+(1-t) F\left(y_{2}\right)
\end{aligned}
$$

so that $F$ is concave. Moreover, $F$ is upper semicontinuous: let $\left(y_{n}\right) \subset L$ be convergent to $y \in L$. For every $x \in K$ and every $n \in \mathbb{N}$ one has $F\left(y_{n}\right) \leq f\left(x, y_{n}\right)$, and taking the limes superior on both sides, we obtain, by continuity of $f$,

$$
\limsup _{n \rightarrow \infty} F\left(y_{n}\right) \leq \limsup _{n \rightarrow \infty} f\left(x, y_{n}\right)=f(x, y)
$$

Since $x \in K$ was arbitrary, this inequality implies $\limsup _{n \rightarrow \infty} F\left(y_{n}\right) \leq F(y)$, i.e. $F$ is upper semicontinuous.

By Theorem E. 38 (applied to $-F$; use also Remark E.39), there exists $\bar{y} \in L$ such that

$$
f(\Phi(\bar{y}), \bar{y})=F(\bar{y})=\sup _{y \in L} F(y) .
$$

We put $\bar{x}=\Phi(\bar{y})$ and show that $(\bar{x}, \bar{y})$ is a saddle point. Clearly, for every $x \in K$,

$$
\begin{equation*}
f(\bar{x}, \bar{y}) \leq f(x, \bar{y}) \tag{E.12}
\end{equation*}
$$

Therefore it remains to show that for every $y \in L$,

$$
\begin{equation*}
f(\bar{x}, \bar{y}) \geq f(\bar{x}, y) \tag{E.13}
\end{equation*}
$$

Let $y \in L$ be arbitrary and put $y_{n}:=\left(1-\frac{1}{n}\right) \bar{y}+\frac{1}{n} y$ and $x_{n}=\Phi\left(y_{n}\right)$. Then, by concavity,

$$
\begin{aligned}
F(\bar{y}) & \geq F\left(y_{n}\right) \quad=\quad f\left(x_{n}, y_{n}\right) \\
& \geq\left(1-\frac{1}{n}\right) f\left(x_{n}, \bar{y}\right)+\frac{1}{n} f\left(x_{n}, y\right) \\
& \geq\left(1-\frac{1}{n}\right) F(\bar{y})+\frac{1}{n} f\left(x_{n}, y\right),
\end{aligned}
$$

or

$$
F(\bar{y}) \geq f\left(x_{n}, y\right) \text { for every } n \in \mathbb{N}
$$

Since $K$ is bounded and closed, the sequence $\left(x_{n}\right) \subset K$ has a weakly convergent subsequence which converges to some element $x_{0} \in K$ (Theorem E. 33 and Corollary E.36). By the preceeding inequality and Corollary E.37,

$$
F(\bar{y}) \geq f\left(x_{0}, y\right)
$$

This is just the remaining inequality (E.13) if we can prove that $x_{0}=\bar{x}$. By concavity, for every $x \in K$ and every $n \in \mathbb{N}$,

$$
\begin{aligned}
f\left(x, y_{n}\right) & \geq f\left(x_{n}, y_{n}\right) \\
& \geq\left(1-\frac{1}{n}\right) f\left(x_{n}, \bar{y}\right)+\frac{1}{n} f\left(x_{n}, y\right) \\
& \geq\left(1-\frac{1}{n}\right) f\left(x_{n}, \bar{y}\right)+\frac{1}{n} F(y) .
\end{aligned}
$$

Letting $n \rightarrow \infty$ in this inequality and using Corollary E. 37 again, we obtain that for every $x \in K$,

$$
f(x, \bar{y}) \geq f\left(x_{0}, \bar{y}\right) .
$$

Hence, $x_{0}=\Phi(\bar{y})=\bar{x}$ and the theorem is proved.

## Appendix $F$ <br> The divergence theorem

by Wolfgang Arendt

## Introduction

This is an appendix to the course "Partial Differential Equations" given in Summer Semester 2006. We give a proof of the Divergence Theorem. The idea is to use Riesz' Theorem to prove uniqueness of the surface measure, i.e., a measure on the boundary of a $C^{1}$-domain $\Omega \subset \mathbb{R}^{n}$ such that

$$
\int_{\Omega} D_{j} v \mathrm{dx}=\int_{\Gamma} v v_{j} d \sigma
$$

for all $j=1, \ldots n$ and all $v \in C^{1}(\bar{\Omega})$, where $v \in C\left(\Gamma, \mathbb{R}^{n}\right)$ denotes the outer normal. In the construction of the surface measure, this uniqueness results makes superfluous the tedious proof of independance of the chosen graph and the partition of the unity.

## F. 1 Open sets of class $C^{1}$

Let $\Omega \subset \mathbb{R}^{n}$ be an open, bounded set and $\Gamma=\partial \Omega$ its boundary. For $x, y \in \mathbb{R}^{n}$ we denote by

$$
x \cdot y=\sum_{j=1}^{n} x_{j} y_{j}
$$

the scalar product, by $|x|:=\sqrt{x \cdot x}$ the Eucleadian norm and by $|x|_{\infty}:=\max _{j=1 \ldots n}\left|x_{j}\right|$ the supremum norm.

Definition F.1. a) Let $U \subset \mathbb{R}^{n}$ be open. We say that $\Gamma \cap U$ is a normal $C^{1}$-graph (with $\Omega$ on one side), if there exist $g \in C^{1}\left(\mathbb{R}^{n-1}\right), r>0, h>0$ such that

$$
U=\left\{(y, g(y)+s): y \in \mathbb{R}^{n-1},|y|_{\infty}<r, s \in \mathbb{R},|s|<h\right\}
$$

such that for $x=(y, g(y)+s) \in U$ one has

$$
\begin{aligned}
& x \in \Omega \text { if and only if } s>0 \\
& x \in \Gamma \text { if and only if } s=0 \text { and } \\
& x \notin \bar{\Omega} \text { if and only if } s<0 .
\end{aligned}
$$

b) Let $U \subset \mathbb{R}^{n}$ be open. We say that $U \cap \Gamma$ is a $C^{1}$-graph with $\Omega$ on one side if there exist an orthonormal $n \times n$ matrix $B$ and $b \in \mathbb{R}^{n}$ such that $\phi(U) \cap \phi(\Gamma)$ is a normal $C^{1}$-graph with $\phi(\Omega)$ on one side where $\phi(x)=B x+b$.
c) We say that $\Omega$ is of class $C^{1}$ or that $\Omega$ has $C^{1}$-boundary if for each $z \in \Gamma$ there exists an open neighborhood $U \subset \mathbb{R}^{n}$ of $z$ such that $U \cap \Gamma$ is a $C^{1}$-graph with $\Omega$ on one side.

Remark F.2. Similarly, we define a normal $C^{k}$-graph or a normal Lipschitz graph, if the function $g$ in a) is of class $C^{k}$ or Lipschitz continuous, respectively. Accordingly, we define $C^{k}$-graphs and Lipschitz graphs as in b ). We say that $\Omega$ is of class $C^{k}$ or $\Omega$ has Lipschitz boundary, if for each $z \in \Gamma$ there exists an open neighborhood $U$ such that $U \cap \Gamma$ is a $C^{k}$-graph or a Lipschitz-graph, respectively.

Next we give the definition of the tangent space of $\Gamma$ at a point $z \in \Gamma$. Here we need this notion merely to define the outer normal intrinsically (Theorem F. 5 below).

Definition F.3. Let $z \in \Gamma$. A vector $v \in \mathbb{R}^{n}$ is said to be tangent to $\Gamma$ at $z$ if there exists $\psi \in C^{1}\left((-\varepsilon, \varepsilon), \mathbb{R}^{n}\right)$ such that

$$
\psi(t) \in \Gamma \quad \text { for } \quad|t|<\varepsilon \quad \text { and } \quad \psi(0)=z, \psi^{\prime}(0)=v
$$

We denote by $T_{z}$ the set of all vectors $v$ which are tangent to $\Gamma$ at $z$. It is easy to see that $T_{z}$ is a vector subspace of $\mathbb{R}^{n}$. We call it the tangent space of $\Gamma$ at $z$.

Proposition F.4. Let $\Omega \subseteq \mathbb{R}^{n}$ be of class $C^{1}$, and let $z \in \Gamma$. Then $T_{z}$ has dimension $n-1$.

Proof. Let $U \subset \mathbb{R}^{n}$ be an open neighborhood of $z$ such that $U \cap \Gamma$ is a $C^{1}$-graph with $\Omega$ on one side. We may assume that the graph is normal and keep the notations of Definition F.1.a). For $x=\left(x_{1}, \ldots x_{n}\right) \in \mathbb{R}^{n}$ we let $\underline{x}:=\left(x_{1}, \ldots x_{n-1}\right) \in \mathbb{R}^{n-1}$. Thus $z_{n}=g(\underline{x})$. We show that

$$
\begin{equation*}
T_{z}=\left\{v \in \mathbb{R}^{n}: v_{n}=\sum_{j=1}^{n-1} v_{j} D_{j} g(\underline{x})\right\} \tag{F.1}
\end{equation*}
$$

which implies the claim. Let $v \in \mathbb{R}^{n}$ such that $v_{n}=\sum_{j=1}^{n-1} v_{j} D_{j} g(\underline{z})$. Define $\psi(t)=$ $(\underline{z})+t \underline{v}, g(\underline{z}+t \underline{v}))$. Then $\psi \in C^{1}\left((-\varepsilon, \varepsilon), \mathbb{R}^{n}\right)$ and $\psi(t) \in \Gamma$ for $|t|<\varepsilon$ if $\varepsilon>0$ is small enough. Then $\psi(0)=z$ and $\psi^{\prime}(0)=\left(\underline{v}, \sum_{j=1}^{n-1} v_{j} D_{j} g(\underline{z})\right)=v$. This proves one inclusion. Conversely, let $v \in T_{z}$. Consider $\psi$ as in Definition F.3. Then $\psi(0)=z, \psi^{\prime}(0)=v$. Since $\psi(t)=\left(\psi_{1}(t), \ldots, \psi_{n}(t)\right) \in \Gamma$ for $|t|<\varepsilon$, it follows that $\psi_{n-1}(t)=g\left(\psi_{1}(t), \ldots, \psi_{n}(t)\right)$ for $|t|<\varepsilon$. Consequently, $v_{n}=\psi_{n}^{\prime}(0)=$ $\sum_{j=1}^{n-1} \psi_{j}^{\prime}(0) D_{j} g\left(\psi_{1}(0), \ldots, \psi_{n-1}(0)\right)=\sum_{j=1}^{n-1} v_{j} D_{j} g(\underline{z})$. This proves the other inclusion.

For $A \subset \mathbb{R}^{n}$ we denote by

$$
A^{\perp}:=\left\{v \in \mathbb{R}^{n}: x \cdot v=0 \quad \text { for all } \quad x \in A\right\}
$$

the space which is orthogonal to $A$. For $x=\left(x_{1}, \ldots, x_{n}\right) \in \mathbb{R}^{n}$ we let $\underline{x}:=$ $\left(x_{1}, \ldots, x_{n-1}\right) \in \mathbb{R}^{n-1}$.
Theorem F. 5 (Definition of the outer normal). Assume that $\Omega$ is of class $C^{1}$. Then for each $z \in \Gamma$ there exists a unique vector $v(z) \in \mathbb{R}^{n}$ satisfying
a) $v(z) \in T_{z}^{\perp},|v(z)|=1$,
b) $z+t v(z) \in \Omega$ for $t \in(-\varepsilon, 0)$,
c) $z+t v(z) \notin \Omega$ for $t \in[0, \varepsilon)$
for some $\varepsilon>0$. We call $v(z)$ the outer normal of $\Omega$ at $z$. Moreover, $v \in C\left(\Gamma, \mathbb{R}^{n}\right)$.
Proof. Since $d T_{z}=n-1$ it follows that for each $z \in \Gamma$ there exists at most one $v(z) \in \mathbb{R}^{n}$ satisfying (a), (b), (c). Let $U \subset \mathbb{R}^{n}$ be open such that $U \cap \Gamma$ is a $C^{1}$-graph. We show that for $z \in \Gamma \cap U$ there exists $v(z)$ satisfying (a), (b), (c). We may assume that the graph is normal and keep the notation of Definition F.1. For $x \in \mathbb{R}^{n}$ we let $\underline{x}=\left(x_{1}, \ldots x_{n-1}\right) \in \mathbb{R}^{n-1}$. Let

$$
\begin{equation*}
v(z)=\frac{(\nabla g(\underline{z}),-1)}{\sqrt{1+\mid \nabla g(\underline{z})^{2}}} \quad \text { for } \quad z \in \Gamma \cap U \tag{F.2}
\end{equation*}
$$

Then $|v(z)|=1$ and $v(z) \in T_{z}^{\perp}$ by (F.1). It remains to show (b) and (c). Let $\psi(x)=$ $g(\underline{x})-x_{n}$. Then $\psi \in C^{1}\left(\mathbb{R}^{n}\right)$ and for $x \in U$ one has

$$
\begin{aligned}
& \psi(x)<0 \text { if and only if } x \in \Omega \quad \text { and } \\
& \psi(x) \geq 0 \text { if and only if } x \notin \Omega
\end{aligned}
$$

Let $z \in \Gamma \cap U$. By Taylor's formula,

$$
\psi(z+h)=(\nabla \psi(z) \mid h)+o(h)
$$

with $\frac{o(h)}{|h|} \rightarrow 0$ as $|h| \rightarrow 0$. Note that $\nabla \psi(z)=(\nabla g(\underline{x}),-1)$, hence $v(z)=\frac{\nabla \psi(z)}{|\nabla \psi(z)|}$. Thus

$$
\psi(z+t v(z))=t(\nabla \psi(z) \cdot v(z))+o(t v(z))=t\left\{|\nabla \psi(z)|+\frac{o(t v(z))}{t}\right\}
$$

Since $|\nabla \psi(z)| \geq 1$ and $\frac{o(t v(z))}{t} \rightarrow 0$ as $t \rightarrow 0$ there exists $\varepsilon>0$ such that $\psi(z+$ $t v(z)) \geq 0$ for $t \in[0, \varepsilon)$ and $\psi(z+t v(z))<0$ for $t \in(-\varepsilon, 0)$. Thus (b) and (c) are proved. It follows from (F.2) that $v: U \cap \Gamma \rightarrow \mathbb{R}^{n}$ is continuous.

## F. 2 The divergence theorem

Let $\Omega \subset \mathbb{R}^{n}$ be a bounded, open set with $C^{1}$-boundary $\Gamma=\partial \Omega$. In this section we introduce the surface measure on $\Gamma$ as the unique measure on $\Gamma$ such that the Divergence Theorem holds. The Divergence Theorem is the extension of the Fundamental Theorem of Calculus to higher dimension. It implies the formula of partial integration in higher dimension and the Green's formulas. The proof of the Divergence Theorem will be given in the next section.

Let $\Omega \subset \mathbb{R}^{n}$ be open and bounded with boundary $\Gamma:=\partial \Omega$. By $C^{1}(\bar{\Omega})$ we denote the space of all continuous functions $u: \bar{\Omega} \rightarrow \mathbb{R}$ which are $C^{1}$ in $\Omega$ such that $D_{j} u$ has a continuous extension to $\bar{\Omega}$ for $j=1, \ldots n$. Here $D_{j} u(x)=\frac{\partial u}{\partial x_{j}}\left(x_{1}, \ldots, x_{n}\right)$ is the $j$ th partial derivative at $x=\left(x_{1}, \ldots, x_{n}\right)$.

Theorem F. 6 (Divergence Theorem). Assume that $\Omega$ is of class $C^{1}$. Then there exists a unique Borel measure $\sigma$ on $\Gamma$ such that

$$
\begin{equation*}
\int_{\Omega} D_{j} u \mathrm{dx}=\int_{\Gamma} u v_{j} d \sigma(j=1, \ldots n) \quad \text { for all } \quad u \in C^{1}(\bar{\Omega}) \tag{F.3}
\end{equation*}
$$

Here $v \in C\left(\Gamma, \mathbb{R}^{n}\right)$ is the outer normal with components

$$
v(z)=\left(v_{1}(z), \ldots, v_{n}(z)\right) \quad(z \in \Gamma)
$$

The measure $\sigma$ is called the surface measure or the $(n-1)$-dimensional Hausdorff measure.

The Divergence Theorem is also called Gauß's Theorem. It is an extension of the Fundamental Theorem of Calculus

$$
\int_{a}^{b} u^{\prime}(x) \mathrm{dx}=u(b)-u(a)
$$

( $u \in C^{1}[a, b]$ ) to higher dimension. The following formula of integration by parts in higher dimension is frequently used.

## Corollary F. 7 (Integration by parts).

$$
\begin{equation*}
\int_{\Omega}\left(D_{j} u\right) v \mathrm{dx}=-\int_{\Omega} u D_{j} v \mathrm{dx}+\int_{\Gamma} u v v_{j} d \sigma \tag{F.4}
\end{equation*}
$$

$\left(u, v \in C^{1}(\bar{\Omega})\right), j=1, \ldots n$, where $v(z)=\left(v_{1}(z), \ldots, v_{n}(z)\right)$ is the outer normal.
Proof. By the Divergence Theorem applied to $u v$ we have

$$
\begin{aligned}
& \int_{\Gamma} u v v_{j} d \sigma=\int_{\Omega} D_{j}(u v) \mathrm{dx} \\
= & \int_{\Omega}\left(D_{j} u\right) v \mathrm{dx}+\int_{\Omega} u\left(D_{j} v\right) \mathrm{dx} .
\end{aligned}
$$

Taking $v=1$ in (F.4) we recover the original formula (F.3). Next we prove the Green's formula. For $u \in C^{1}(\bar{\Omega})$ we denote by

$$
\frac{\partial u}{\partial v}(z):=\sum_{j=1}^{n} D_{j} u(z) v_{j}(z)
$$

the normal derivative. Recall, that for $x \in \Omega, v \in \mathbb{R}^{n}$,

$$
\begin{aligned}
D_{v} u(x) & :=\left.\frac{\partial}{\partial t}\right|_{t=0} u(x+t v) \\
& =\nabla u(x) \cdot v
\end{aligned}
$$

is the directional derivative (or Gateaux derivative) of $u$ at $x$ in the direction $v$. Here $\nabla u(x) \cdot v=\sum_{j=1}^{n} D_{j} u(x) v_{j}$ is the scalar product in $\mathbb{R}^{n}$. Thus

$$
\frac{\partial u}{\partial v}(z)=D_{v(z)} u(z)
$$

if $u$ has an extension in $C^{1}\left(\mathbb{R}^{n}\right)$.
Corollary F. 8 (Green's formula). Let $u, v \in C^{2}(\bar{\Omega})$. Then
a) $\int_{\Omega} \Delta u \cdot v \mathrm{dx}=-\int_{\Omega} \nabla u \nabla v \mathrm{dx}+\int_{\Gamma} \frac{\partial u}{\partial v} v d \sigma$
b) $\int_{\Omega} \Delta u \mathrm{dx}=\int_{\Gamma} \frac{\partial u}{\partial v} d \sigma$.

Here $C^{2}(\bar{\Omega})=\left\{u \in C^{1}(\bar{\Omega}): D_{j} u \in C^{1}(\bar{\Omega}), j=1 \ldots n\right\}$.
Proof. (a) Applying the integration-by-parts formula (F.4) to $D_{j} u$ instead of $u$ we obtain

$$
\begin{aligned}
\int_{\Omega} \Delta u v \mathrm{dx} & =\sum_{j=1}^{n} \int_{\Omega} D_{j}\left(D_{j} u\right) v \mathrm{dx} \\
& =\sum_{j=1}^{n}\left\{-\int_{\Omega} D_{j} u D_{j} v \mathrm{dx}+\int_{\Gamma} v_{j} D_{j} u v d \sigma\right\} \\
& =-\int_{\Omega} \nabla u \cdot \nabla v \mathrm{dx}+\int_{\Gamma} \frac{\partial u}{\partial v} v d \sigma
\end{aligned}
$$

(b) follows from (a) taking $v=1$.

Exercise. For $u, v \in C^{2}(\bar{\Omega})$,

$$
\int_{\Omega}((\Delta u) v-u \Delta v) \mathrm{dx}=\int_{\Gamma}\left(\frac{\partial u}{\partial v} v-u \frac{\partial v}{\partial v}\right) d \sigma
$$

Finally we give some comments explaining the name "Divergence Theorem" and also the physical significance of this result. Let $v \in C^{1}\left(\bar{\Omega}, \mathbb{R}^{n}\right)$, be a vector field, $v(x)=\left(v_{1}(x), \ldots, v_{n}(x)\right)$ with $v_{j} \in C^{1}(\bar{\Omega}), j=1, \ldots n$. Denote by

$$
(\operatorname{div} v)(x):=\sum_{j=1}^{n} D_{j} v_{j}(x)
$$

the divergence of $v$. Note that $\operatorname{div} v \in C(\bar{\Omega})$. $\operatorname{By}(v(z) \cdot v(z))=\sum_{j=1}^{n} v_{j}(z) v_{j}(z) \quad(z \in$ $\Gamma)$ we denote the scalar product of $v$ and the normal vector on $\Gamma$.

Corollary F.9. Let $v \in C^{1}\left(\bar{\Omega}, \mathbb{R}^{n}\right)$. Then

$$
\begin{equation*}
\int_{\Omega}(\operatorname{div} v)(x) \mathrm{dx}=\int_{\Gamma} v \cdot v d \sigma \tag{F.5}
\end{equation*}
$$

Proof. One has

$$
\int_{\Omega}(\operatorname{div} v)(x) \mathrm{dx}=\sum_{j=1}^{n} \int_{\Omega} D_{j} v_{j}(x) \mathrm{dx}=\sum_{j=1}^{n} \int_{\Gamma} v_{j}(z) v_{j}(z) d \sigma(z)=\int_{\Gamma} v \cdot v d \sigma
$$

by Theorem F.6.
Remark F. 10 (Physical interpretation). The scalar product $v(z) \cdot v(z)=$ $\|v(z)\| \cos \alpha$, where $\alpha$ is the angle between the vectors $v(z)$ and $v(z)$, is the projection of $v(z)$ to the normal vector $v(z)$. If $d \sigma(z)$ is identified with a small square (we consider dimension $n=3$ ), then $v(z) \cdot v(z) d \sigma(z)$ is the flux of the vectorfield $v$ through the surface element $d \sigma(z)$. Thus, if $v$ is the velocity of a liquid, measured in $m / \mathrm{sec}$, then $v(z) \cdot v(z) d \sigma(z)$ is the quantity of liquid going through the surface element $d \sigma(z)$ in 1 second. The integral
F. 3 Proof of the divergence theorem

$$
\begin{equation*}
\int_{\partial \omega} v(z) \cdot v(z) d \sigma(z) \tag{F.6}
\end{equation*}
$$

is the quantity of liquid going through the surface of a body $\omega$ in 1 second. If the liquid is incompressible, then the quantity going into the body is equal to the quantity going out of it; hence the flux (F.5) is 0 . By the Divergence Theorem in the form of Corollary F. 9 this implies that

$$
\int_{\omega}(\operatorname{div} v)(x) \mathrm{dx}=0
$$

for all open sets $\omega$ of class $C^{1}$ satisfying $\bar{\omega} \subset \Omega$. This implies that

$$
\operatorname{div} v(x)=0 \quad \text { for all } \quad x \in \Omega .
$$

This result is rephrased by saying that the velocity field of an incompressible liquid is divergence free. The same is true for the electrical field in a space without charge.

Example F. 11 (Archimedean Principle). A body $A$ is in a liquid of constant density $c>0$, whose surface coincides with the plane $\left\{\left(x_{1}, x_{2}, 0\right): x_{1}, x_{2} \in \mathbb{R}\right\}$. At the point $z \in \partial A$ the liquid yields a pressure onto the body $A$ which is equal to $c z_{3} v(z)$. Observe that $z_{3}$ is negative and the presure is directed to the interior of $A$. Thus, the force $F$ executed onto the body $A$ is

$$
F=\int_{\partial A} c z_{3} v(z) d \sigma(z)
$$

i.e., $F_{i}=\int_{\partial A} c z_{3} v_{i}(z) d \sigma(z) \quad(i=1,2,3)$. By the Divergence Theorem,

$$
F_{i}=\int_{A} c \frac{\partial}{\partial x_{i}} x_{3} \mathrm{dx}
$$

Hence $F_{1}=F_{2}=0$ and

$$
F_{3}=c \int_{A} \mathrm{dx}=c|A|
$$

where $|A|$ is the volume of $A$.

## F. 3 Proof of the divergence theorem

We start recalling Riesz's Theorem. Let $K$ be a compact space and $C(K)$ the space of all continuous functions $f: K \rightarrow \mathbb{R}$. By a positive linear form $\varphi$ on $C(K)$ we understand a linear mapping $\varphi: C(K) \rightarrow \mathbb{R}$ such that

$$
f \geq 0 \quad \text { implies } \quad \varphi(f) \geq 0
$$

for all $f \in C(K)$. If $\mu$ is a Borel measure on $K$, then

$$
\varphi(f)=\int_{K} f(x) d \mu(x)
$$

defines a positive linear form on $C(K)$. Conversely, Riesz's Theorem says that each positive linear form is induced by a measure.

Theorem F. 12 (Riesz). Let $\varphi$ be a positive linear form on $C(K)$. Then there exists a unique Borel measure $\mu$ on $K$ such that

$$
\begin{equation*}
\varphi(f)=\int_{K} f(x) d \mu(x) \tag{F.7}
\end{equation*}
$$

for all $f \in C(K)$.
Next we recall the Stone-Weierstraß Theorem. A subalgebra of $C(K)$ is a subspace $\mathscr{A}$ of $C(K)$ such that

$$
f, g \in \mathscr{A} \quad \text { implies } \quad f \cdot g \in \mathscr{A}
$$

We say that $\mathscr{A}$ separates the points of $K$ if for all $x, y \in K$ satisfying $x \neq y$ there exists $f \in \mathscr{A}$ such that $f(x) \neq f(y)$.

Theorem F. 13 (Stone-Weierstraß). Let $\mathscr{A}$ be a subalgebra of $C(K)$ which separates the points of $K$ and contains the constant-1-function $1_{K}$. Then $\mathscr{A}$ is dense in $C(K)$.

Now let $\Omega \subset \mathbb{R}^{n}$ be an open, bounded set of class $C^{1}$. We first prove uniqueness of the surface measure.

Proposition F. 14 (Uniqueness of the surface measure). Let $\sigma$ be a Borel measure on $\Gamma$ such that

$$
\begin{equation*}
\int_{\Gamma} u v_{j} d \sigma=0 \quad(j=1, \ldots, n) \tag{F.8}
\end{equation*}
$$

for all $u \in C^{1}(\bar{\Omega})$. Then $\sigma=0$.
Proof. The set $\mathscr{A}=\left\{u_{\mid \Gamma}: u \in C^{1}\left(\mathbb{R}^{n}\right)\right\}$ is dense in $C(\Gamma)$ by the StoneWeierstraß Theorem. It follows from this and the assumption (F.8) that $\int_{\Gamma} u v_{j} d \sigma=0$ for all $u \in C(\Gamma)$. Replacing $u$ by $u v_{j}$ we conclude that $\int_{\Gamma} u v_{j}^{2} d \sigma=0$ for all $u \in C(\Gamma)$. Since $\sum_{j=1}^{n} v_{j}^{2}=1$, it follows that $\int_{\Gamma} u d \sigma=0$. This implies that $\sigma=0$, by Riesz' Theorem.

Next we prove existence of the surface measure. Let $U \subset \mathbb{R}^{n}$ be open such that $U \cap \Gamma \neq \emptyset$. Let

$$
C_{c}(U \cap \Gamma):=\{u \in C(\Gamma): \exists K \subset U \text { compact such that } u(z)=0 \text { for } z \in \Gamma \backslash K\}
$$

We denote by $C_{c}(U \cap \Gamma)_{+}^{\prime}:=\left\{\varphi: C_{c}(U \cap \Gamma) \rightarrow \mathbb{R}: \varphi\right.$ is linear such that $f \geq 0$ implies $\varphi(f) \geq 0$ for all $\left.f \in C_{c}(U \cap \Gamma)\right\}$ the set of all positive linear forms on $C_{c}(U \cap \Gamma)$. Let $U \cap \Gamma$ be a normal graph. Define

$$
\begin{equation*}
\varphi(f)=\int_{|y|_{\infty}<r} f(y, g(y)) \sqrt{1+|\nabla g(y)|^{2}} \mathrm{dy} \tag{F.9}
\end{equation*}
$$

for all $f \in C_{c}(U \cap \Gamma)$ where we use the notation of Definition F.1. Then $\varphi \in C_{c}(U \cap$ $\Gamma)_{+}^{\prime}$.

Lemma F.15. Let $U \cap \Gamma$ be a normal $C^{1}$-graph as above. Let $u \in C^{1}(\bar{\Omega})$ be such that $\operatorname{supp} u \subset U$. Then

$$
\begin{equation*}
\int_{\Omega} D_{j} u \mathrm{dx}=\varphi\left(u_{\left.\right|_{\Gamma}} v_{j}\right) \quad(j=1 \ldots n) \tag{F.10}
\end{equation*}
$$

Here supp $u:=\overline{\{x \in \Omega: u(x) \neq 0\}}$ (i.e. the closure of the open set $\{x \in \Omega$ : $u(x) \neq 0\})$ is a compact subset of $\mathbb{R}^{n}$. We assume that supp $u \subset U$, hence $u(x)=0$ for all $x \in U \backslash \operatorname{supp} u$. Thus $u_{\left.\right|_{\Gamma}} \in C_{c}(U \cap \Gamma)$.

Proof. Denote by

$$
\phi: \mathbb{R}^{n-1} \times(-h, h) \rightarrow U
$$

the diffeomorphism given by $\phi(y, s)=(y, g(y)+s)$. Then $\operatorname{det} D \phi(y, s)=1$ for all $(y, s) \in \mathbb{R}^{n-1} \times(-h, h)$.
a) Let $j \in\{1, \ldots, n-1\}$. Let $u \in C^{1}(\bar{\Omega})$ such that $\operatorname{supp} u \subset U$. Then

$$
\begin{aligned}
& \int_{\Omega} D_{j} u \mathrm{dx}=\int_{|y|_{\infty}<r} \int_{0}^{h} D_{j} u(y, g(y)+s) \mathrm{dy} \\
= & \int_{|y|_{\infty}<r} \int_{0}^{h} \frac{\partial}{\partial j} u(y, g(y)+s) \mathrm{dy}-\int_{|y|_{\infty}<r} \int_{0}^{h} D_{n} u(y, g(y)+s) D_{j} g(y) \mathrm{dy} \\
= & \int_{0}^{h} \int_{|y|_{\infty}<r} \frac{\partial}{\partial y_{j}} u(y, g(y)+s) \mathrm{dy}+\int_{|y|_{\infty}<r} u(y, g(y)) D_{j} g(y) \mathrm{dy} \\
= & \int_{|y|_{\infty}<r} u(y, g(y)) D_{j} g(y) \mathrm{dy}
\end{aligned}
$$

where we used the Fundamental Theorem of Calculus in 1 variable and the fact that $\operatorname{supp} u \subset U$. Since $v_{j}(y, g(y))=\frac{D_{j} g(y)}{\sqrt{1+|\nabla g(y)|^{2}}}$, the last term equals $\varphi\left(v_{j} u_{\left.\right|_{\Gamma}}\right)$.
b) Let $j=n$. Then $\frac{\partial u}{\partial s}(y, g(y)+s)=D_{n} u(y, g(y)+s)$. Hence

$$
\begin{aligned}
\int_{\Omega} D_{n} u(x) \mathrm{dx} & =\int_{|y|_{\infty}<r} \int_{0}^{h} D_{n} u(y, g(y)+s) \mathrm{dy}= \\
\int_{|y|_{\infty}<r} \int_{0}^{h} \frac{\partial}{\partial s} u(y, g(y)+s) \mathrm{dy} & =-\int_{|y|_{\infty}<r} u(y, g(y)) \mathrm{dy}= \\
\int_{|y|_{\infty}<r} u(y, g(y)) v_{n}(y, g(y)) \sqrt{1+|\nabla g(y)|^{2}} \mathrm{dy} & =\varphi\left(u_{\left.\right|_{\Gamma}} v_{n}\right)
\end{aligned}
$$

since $v_{n}(y, g(y))=\frac{-1}{\sqrt{1+|\nabla g(y)|^{2}}}$.
Now let $U \subset \mathbb{R}^{n}$ be open and suppose that $U \cap \Gamma$ is a $C^{1}$-graph, i.e. there exist an orthogonal matrix $B$, det $B=1$, and $b \in \mathbb{R}^{n}$ such that $\tilde{\Gamma} \cap \tilde{U}$ is a normal $C^{1}$-graph with $\tilde{\Omega}$ on one side where $\phi(x)=B x+b, \tilde{\Omega}=\phi(\Omega), \tilde{\Gamma}=\partial \tilde{\Omega}=\phi(\Gamma), \tilde{U}=\phi(U)$. Notice that the outer normal $\tilde{v}(\phi(z))$ of $\tilde{\Omega}$ at $\phi(z)$ is given by

$$
\begin{equation*}
\tilde{v}(\phi(z))=B v(z) \tag{F.11}
\end{equation*}
$$

for all $z \in \Gamma$. Consider the positive linear form $\tilde{\varphi}$ on $C_{c}(\tilde{\Gamma} \cap \tilde{U})$ constructed above. Define $\varphi \in C_{c}(\Gamma \cap U)_{+}^{\prime}$ by

$$
\begin{equation*}
\varphi(f)=\tilde{\varphi}\left(f \circ \phi^{-1}\right) \tag{F.12}
\end{equation*}
$$

for $f \in C_{c}(\Gamma \cap U)$.
Lemma F.16. Let $U \cap \Gamma$ be a $C^{1}$-graph as above. Let $u \in C^{1}(\bar{\Omega})$ such that $\operatorname{supp} u \subset$ $U$. Then

$$
\int_{\Omega} D_{j} u \mathrm{dx}=\varphi\left(u_{\mid \Gamma} v_{j}\right) \quad(j=1, \ldots, n)
$$

Proof. Let $B=\left(b_{k j}\right)_{k, j=1 \ldots n}$. Let $\tilde{u}=u \circ \phi^{-1}$. Then $\tilde{u} \in C^{1}(\bar{\Omega})$ and by Lemma F. 15 .

$$
\begin{aligned}
\int_{\Omega} D_{j} u & =\int_{\Omega} \frac{\partial}{\partial x_{j}} \tilde{u}(\phi(x)) \mathrm{dx} \\
& =\int_{\Omega} \sum_{k=1}^{n} D_{k} \tilde{u}(\phi(x)) b_{k j} \mathrm{dx} \\
& =\int_{\tilde{\Omega}} \sum_{k=1}^{n} D_{k} \tilde{u}(y) b_{k j} \mathrm{dy} \\
& =\sum_{k=1}^{n} b_{k j} \tilde{\varphi}\left(\tilde{u}_{\mid \tilde{\Gamma}} \tilde{v}_{k}\right) \\
& =\tilde{\varphi}\left(u \circ \phi_{\mid \tilde{\Gamma}}^{-1}\left(B^{-1} \tilde{v}\right)_{j}\right) \\
& =\varphi\left(u_{\mid \Gamma}\left(B^{-1} \tilde{v} \circ \phi\right)_{j}\right) \\
& =\varphi\left(u_{\mid \Gamma} v_{j}\right) .
\end{aligned}
$$

Now we piece together the surface measure using the positive linear forms on the graphs. Since $\Omega$ is of class $C^{1}$ there exist open sets $U_{m} \subset \mathbb{R}^{n}, m=1, \ldots, M$ such that

$$
\Gamma \subset \bigcup_{m=1}^{M} U_{m}
$$

and $\Gamma \cap U_{m}$ is a $C^{1}$-graph for each $m \in\{1, \ldots, M\}$. Let $U_{0}$ be an open set such that $\bar{\Omega} \subset \bigcup_{m=0}^{M} U_{m}$. We recall that there exists a partition of unity on $\bar{\Omega}$ subordinate to the open sets $U_{m}, m=0,1 \ldots M$; that is, there exist functions $\eta_{m} \in C^{\infty}\left(\mathbb{R}^{n}\right)$ such that $0 \leq \eta_{m} \leq 1$, supp $\eta_{m} \subset U_{m}$ and

$$
\sum_{m=0}^{M} \eta_{m}(x)=1
$$

for all $x \in \bar{\Omega}$. For $m \in\{1, \ldots, M\}$ let $\varphi_{m} \in C_{c}\left(\Gamma \cap U_{m}\right)_{+}^{\prime}$ be the positive linear form constructed above; that is, $\varphi_{m}$ satisfies

$$
\int_{\Omega} D_{j} u=\varphi_{m}\left(v_{j} u_{\mid\left\lceil\cap U_{m}\right.}\right)
$$

for all $u \in C^{1}(\bar{\Omega})$ satisfying supp $u \subset U_{m}$. Define the positive linear form $\varphi$ on $C(\Gamma)$ by

$$
\varphi(f):=\sum_{m=1}^{M} \varphi_{m}\left(\eta_{m} f\right)
$$

By Riesz' Theorem there exists a Borel measure $\sigma$ on $\Gamma$ such that

$$
\varphi(f)=\int_{\Gamma} f(z) d \sigma(z)
$$

for all $f \in C(\Gamma)$. Let $u \in C^{1}(\bar{\Omega})$ and let $j \in\{1, \ldots n\}$. Then $u \eta_{0}$ has compact support in $\Omega$. Hence $\int_{\Omega} D_{j}\left(u \eta_{0}\right) \mathrm{dx}=0$. Indeed, extending $u \eta_{0}$ by 0 outside of $\Omega$ we obtain a function $v \in C^{1}\left(\mathbb{R}^{n}\right)$ vanishing outside of a cube $Q:=\left\{x:\left|x_{j}\right| \leq R\right.$ for all $j=$ $1, \ldots, n\}$. Since

$$
\int_{-R}^{R} \frac{\partial v}{\partial x_{j}}\left(x_{1} \ldots x_{j}, x \ldots x_{n}\right) \mathrm{dx} \mathrm{x}_{j}=v\left(x_{1}, \ldots, R, \ldots x_{n}\right)-v\left(x_{1}, \ldots,-R, \ldots x_{n}\right)=0
$$

for all $x_{1}, \ldots x_{j-1}, x_{j+1}, \ldots x_{n} \in \mathbb{R}$ it follows that

$$
\begin{aligned}
& \int_{\Omega} D_{j}\left(u \eta_{0}\right) \mathrm{dx}=\int_{Q} D_{j} v \mathrm{dx} \\
& =\int_{-R}^{R} \ldots \int_{-R}^{R} \frac{\partial v}{\partial x_{j}}\left(x_{1} \ldots x_{j} \ldots x_{n}\right) \mathrm{dx}_{j} \mathrm{dx}_{1} \ldots \mathrm{dx}_{j-1} \mathrm{dx}_{j+1} \ldots \mathrm{dx}_{1}=0 .
\end{aligned}
$$

Let $m \in\{1 \ldots M\}$. Then $u \eta_{m} \in C_{c}\left(\Gamma \cap U_{j}\right)$. Hence for $j \in\{1, \ldots n\}$,

$$
\int_{\Omega} D_{j}\left(\eta_{m} u\right) \mathrm{dx}=\varphi_{m}\left(\eta_{m} u_{\mid \Gamma} v_{j}\right)
$$

Thus

$$
\begin{aligned}
\int_{\Omega} D_{j} u \mathrm{dx} & =\int_{\Omega} D_{j}\left(\sum_{m=0}^{M} \eta_{m} u\right) \mathrm{dx} \\
& =\sum_{m=0}^{M} \int_{\Omega^{\prime}} D_{j}\left(\eta_{m} u\right) \mathrm{dx} \\
& =\sum_{m=1}^{M} \int_{\Omega^{2}} D_{j}\left(\eta_{m} u\right) \mathrm{dx} \\
& =\sum_{m=1}^{M} \varphi_{m}\left(\eta_{m} u_{\mid \Gamma} v_{j}\right) \\
& =\varphi\left(u_{\mid \Gamma} v_{j}\right) \\
& =\int_{\Gamma} u_{\left.\right|_{\Gamma}} v_{j} d \sigma
\end{aligned}
$$

Thus $\sigma$ is a Borel measure satisfying the requirements of the Divergence Theorem. We have seen before that it is unique. In particular, the construction given above does not depend on the choice of the graphs and the choice of the partition of unity.

## Appendix G

## Sobolev spaces

## G. 1 Test functions, convolution and regularization

Let $\Omega \subseteq \mathbb{R}^{d}$ be an open set. For every continuous function $\varphi \in C(\Omega)$ we define the support

$$
\operatorname{supp} \varphi:=\overline{\{x \in \Omega: \varphi(x) \neq 0\}}
$$

where the closure is to be understood in $\mathbb{R}^{d}$. Thus, the support is by definition always closed in $\mathbb{R}^{d}$, but it is not necessarily a subset of $\Omega$. Next we let

$$
\mathscr{D}(\Omega):=C_{c}^{\infty}(\Omega):=\left\{\varphi \in C^{\infty}(\Omega): \operatorname{supp} \varphi \subset \Omega \text { is compact }\right\}
$$

be the space of test functions on $\Omega$, and

$$
L_{l o c}^{1}(\Omega):=\left\{f: \Omega \rightarrow \mathbb{K} \text { measurable }: \int_{K}|f|<\infty \forall K \subset \Omega \text { compact }\right\}
$$

the space of locally integrable functions on $\Omega$.
For every $f \in L_{l o c}^{1}\left(\mathbb{R}^{d}\right)$ and every $\varphi \in \mathscr{D}\left(\mathbb{R}^{d}\right)$ we define the convolution $f * \varphi$ by

$$
\begin{aligned}
f * \varphi(x) & :=\int_{\mathbb{R}^{d}} f(x-y) \varphi(y) d y \\
& =\int_{\mathbb{R}^{d}} f(y) \varphi(x-y) d y
\end{aligned}
$$

Lemma G.1. For every $f \in L_{\text {loc }}^{1}\left(\mathbb{R}^{d}\right)$ and every $\varphi \in \mathscr{D}\left(\mathbb{R}^{d}\right)$ one has $f * \varphi \in C^{\infty}\left(\mathbb{R}^{d}\right)$ and for every $1 \leq i \leq d$,

$$
\frac{\partial}{\partial x_{i}}(f * \varphi)=f * \frac{\partial \varphi}{\partial x_{i}}
$$

Proof. Let $e_{i} \in \mathbb{R}^{d}$ be the $i$-th unit vector. Then

$$
\lim _{h \rightarrow 0} \frac{1}{h}\left(\varphi\left(x+h e_{i}\right)-\varphi(x)\right)=\frac{\partial \varphi}{\partial x_{i}}(x)
$$

uniformly in $x \in \mathbb{R}^{d}$ (note that $\varphi$ has compact support). Hence, for every $x \in \mathbb{R}^{d}$

$$
\begin{aligned}
& \frac{1}{h}\left(f * \varphi\left(x+h e_{i}\right)-f * \varphi(x)\right) \\
= & \frac{1}{h} \int_{\mathbb{R}^{d}} f(y)\left(\varphi\left(x+h e_{i}-y\right)-\varphi(x-y)\right) d y \\
\rightarrow & \int_{\mathbb{R}^{d}} f(y) \frac{\partial \varphi}{\partial x_{i}}(x-y) d y .
\end{aligned}
$$

The following theorem is proved in courses on measure theory. We omit the proof.

Theorem G. 2 (Young's inequality). Let $f \in L^{p}\left(\mathbb{R}^{d}\right)$ and $\varphi \in \mathscr{D}\left(\mathbb{R}^{d}\right)$. Then $f * \varphi \in$ $L^{p}\left(\mathbb{R}^{d}\right)$ and

$$
\|f * \varphi\|_{p} \leq\|f\|_{p}\|\varphi\|_{1}
$$

Theorem G.3. For every $1 \leq p<\infty$ and every open $\Omega \subset \mathbb{R}^{d}$ the space $\mathscr{D}(\Omega)$ is dense in $L^{p}(\Omega)$.

Proof. The technique of this proof (regularization and truncation) is important in the theory of partial differential equations, distributions and Sobolev spaces. The first step (regularization) is based on Lemma G.1. The truncation step is in this case relatively easy.

Regularization. Let $\varphi \in \mathscr{D}\left(\mathbb{R}^{d}\right)$ be a positive function such that $\|\varphi\|_{1}=\int_{\mathbb{R}^{d}} \varphi=$ 1. One may take for example the function

$$
\varphi(x):= \begin{cases}c e^{1 /\left(1-|x|^{2}\right)} & \text { if }|x|<1  \tag{G.1}\\ 0 & \text { otherwise }\end{cases}
$$

with an appropriate constant $c>0$. Then let $\varphi_{n}(x):=n^{d} \varphi(n x)$, so that $\left\|\varphi_{n}\right\|_{1}=$ $\int_{\mathbb{R}^{d}} \varphi_{n}=1$ for every $n \in \mathbb{N}$.

Let $f \in L^{p}\left(\mathbb{R}^{d}\right)$. By Lemma G. 1 and Young's inequality (Theorem G.2), for every $n \in \mathbb{N}, f_{n}:=f * \varphi_{n} \in C^{\infty}\left(\mathbb{R}^{d}\right) \cap L^{p}\left(\mathbb{R}^{d}\right)$ and $\left\|f_{n}\right\|_{p} \leq\|f\|_{p}$. Hence, for every $n \in \mathbb{N}$ the operator $T_{n}: L^{p}\left(\mathbb{R}^{d}\right) \rightarrow L^{p}\left(\mathbb{R}^{d}\right), f \mapsto f * \varphi_{n}$ is linear and bounded and $\left\|T_{n}\right\| \leq 1$. Moreover, if $f=1_{I}$ for some bounded interval $I=\left(a_{1}, b_{1}\right) \times \cdots \times\left(a_{d}, b_{d}\right) \subset \Omega$, then

$$
\begin{aligned}
\left\|f_{n}-f\right\|_{p}^{p} & =\int_{\mathbb{R}^{d}}\left|\int_{\mathbb{R}^{d}} f(x-y) \varphi(n y) n^{d} d y-f(x)\right|^{p} d x \\
& =\int_{\mathbb{R}^{d}}\left|\int_{\mathbb{R}^{d}}\left(f\left(x-\frac{y}{n}\right)-f(x)\right) \varphi(y) d y\right|^{p} d x \\
& \leq \int_{\mathbb{R}^{d}}\left(\int_{\mathbb{R}^{d}}\left|f\left(x-\frac{y}{n}\right)-f(x)\right| \varphi(y) d y\right)^{p} d x \rightarrow 0
\end{aligned}
$$

as $n \rightarrow \infty$ by Lebesgue's dominated convergence theorem. In other words, $\lim _{n \rightarrow \infty}\left\|T_{n} f-f\right\|_{p}=0$ for every $f=1_{I}$ with $I$ as above. Since $\operatorname{span}\left\{1_{I}: I \subset \mathbb{R}^{d}\right.$
bounded interval $\}$ is dense in $L^{p}\left(\mathbb{R}^{d}\right)$, we find that $\lim _{n \rightarrow \infty}\left\|T_{n} f-f\right\|_{p}=0$ for every $f$ from a dense subset $M$ of $L^{p}\left(\mathbb{R}^{d}\right)$. Since the $T_{n}$ are bounded, we conclude from Lemma D. 58 that $T_{n} f \rightarrow f$ in $L^{p}\left(\mathbb{R}^{d}\right)$ for every $f \in L^{p}\left(\mathbb{R}^{d}\right)$. This proves that $L^{p} \cap C^{\infty}\left(\mathbb{R}^{d}\right)$ is dense in $L^{p}\left(\mathbb{R}^{d}\right)$.

Truncation. Now we consider a general open set $\Omega \subset \mathbb{R}^{d}$ and prove the claim. Let $\varphi \in \mathscr{D}\left(\mathbb{R}^{d}\right)$ be a positive test function such that $\operatorname{supp} \varphi \subset \overline{B(0,1)}$ and $\int_{\mathbb{R}^{d}} \varphi=1$ (one may take for example the function from (G.1)). Then let $\varphi_{n}(x):=n^{d} \varphi(n x)$.

For every $n \in \mathbb{N}$ we let

$$
K_{n}:=\left\{x \in \Omega: \operatorname{dist}(x, \partial \Omega) \geq \frac{1}{n}\right\} \cap \overline{B(0, n)},
$$

so that $K_{n} \subset \Omega$ is compact for every $n \in \mathbb{N}$.
Now let $f \in L^{p}(\Omega) \subset L^{p}\left(\mathbb{R}^{d}\right)$ and $\varepsilon>0$. Let

$$
f 1_{K_{n}}(x)= \begin{cases}f(x) & \text { if } x \in K_{n} \\ 0 & \text { if } x \in \Omega \backslash K_{n}\end{cases}
$$

By Lebesgue's dominated convergence theorem (since $\bigcup_{n} K_{n}=\Omega$ ),

$$
\left\|f-f 1_{K_{n}}\right\|_{p}^{p}=\int_{\Omega}|f|^{p}\left(1-1_{K_{n}}\right)^{p} \rightarrow 0 \text { as } n \rightarrow \infty
$$

In particular, there exists $n \in \mathbb{N}$ such that $\left\|f-f 1_{K_{n}}\right\|_{p} \leq \varepsilon$.
For every $m \geq 4 n$ we define $g_{m}:=\left(f 1_{K_{n}}\right) * \varphi_{m} \in L^{p} \cap C^{\infty}\left(\mathbb{R}^{d}\right)$; note that we here consider $L^{p}(\Omega)$ as a subspace of $L^{p}\left(\mathbb{R}^{d}\right)$ by extending functions in $L^{p}(\Omega)$ by 0 outside $\Omega$. However, since $g_{m}=0$ outside $K_{2 n}$, we find that actually $g_{m} \in \mathscr{D}(\Omega)$. By the first step (regularisation), there exists $m \geq 4 n$ so large that $\left\|g_{m}-f 1_{K_{n}}\right\|_{p} \leq \varepsilon$. For such $m$ we have $\left\|f-g_{m}\right\|_{p} \leq 2 \varepsilon$, and the claim is proved.

Lemma G.4. Let $f \in L_{l o c}^{1}(\Omega)$ be such that

$$
\int_{\Omega} f \varphi=0 \quad \text { for every } \varphi \in \mathscr{D}(\Omega) .
$$

Then $f=0$.
Proof. We first assume that $f \in L^{1}(\Omega)$ is real and that $\Omega$ has finite measure. By Theorem G.3, for every $\varepsilon>0$ there exists $g \in \mathscr{D}(\Omega)$ such that $\|f-g\|_{1} \leq \varepsilon$. By assumption, this implies

$$
\left|\int_{\Omega} g \varphi\right|=\left|\int_{\Omega}(f-g) \varphi\right| \leq \varepsilon\|\varphi\|_{\infty} \quad \forall \varphi \in \mathscr{D}(\Omega) .
$$

Let $K_{1}:=\{x \in \Omega: g(x) \geq \varepsilon\}$ and $K_{2}:=\{x \in \Omega: g(x) \leq-\varepsilon\}$. Since $g$ is a test function, the sets $K_{1}, K_{2}$ are compact. Since they are disjoint and do not touch the boundary of $\Omega$,

$$
\inf \left\{|x-y|,|x-z|,|y-z|: x \in K_{1}, y \in K_{2}, z \in \partial \Omega\right\}=: \delta>0
$$

Let $K_{i}^{\delta}:=\left\{x \in \Omega: \operatorname{dist}\left(x, K_{i}\right) \leq \delta / 4\right\}(i=1,2)$. Then $K_{1}^{\delta}$ and $K_{2}^{\delta}$ are two compact disjoint subsets of $\Omega$. Let

$$
h(x):= \begin{cases}1 & \text { if } x \in K_{1}^{\delta} \\ -1 & \text { if } x \in K_{2}^{\delta} \\ 0 & \text { else }\end{cases}
$$

choose a positive test function $\varphi \in \mathscr{D}\left(\mathbb{R}^{d}\right)$ such that $\int_{\mathbb{R}^{d}} \varphi=1$ and $\operatorname{supp} \varphi \subset$ $B(0, \delta / 8)$, and let $\psi:=h * \varphi$. Then $\psi \in \mathscr{D}(\Omega),-1 \leq \psi \leq 1, \psi=1$ on $K_{1}$ and $\psi=-1$ on $K_{2}$. Let $K:=K_{1} \cup K_{2}$. Then

$$
\int_{K}|g|=\int_{K} g \psi \leq \varepsilon+\int_{\Omega \backslash K}|g \psi| \leq \varepsilon+\int_{\Omega \backslash K}|g|
$$

Hence,

$$
\int_{\Omega}|g|=\int_{K}|g|+\int_{\Omega \backslash K}|g| \leq \varepsilon+2 \int_{\Omega \backslash K}|g| \leq \varepsilon(1+2|\Omega|),
$$

which implies

$$
\int_{\Omega}|f| \leq \int_{\Omega}|f-g|+\int_{\Omega}|g| \leq 2 \varepsilon(1+|\Omega|)
$$

Since $\varepsilon>0$ was arbitrary, we find that $f=0$.
The general case can be obtained from the particular case ( $f \in L^{1}$ and $|\Omega|<\infty$ ) by considering first real and imaginary part of $f$ separately, and then by considering $f 1_{B}$ for all closed (compact) balls $B \subset \Omega$.

## G. 2 Sobolev spaces in one dimension

Recall the fundamental rule of partial integration: if $f, g \in C^{1}([a, b])$ on some compact interval $[a, b]$, then

$$
\int_{a}^{b} f g^{\prime}=f(b) g(b)-f(a) g(a)-\int_{a}^{b} f^{\prime} g
$$

In particular, for every $f \in C^{1}([a, b])$ and every $\varphi \in \mathscr{D}(a, b)$

$$
\begin{equation*}
\int_{a}^{b} f \varphi^{\prime}=-\int_{a}^{b} f^{\prime} \varphi \tag{G.2}
\end{equation*}
$$

since $\varphi(a)=\varphi(b)=0$.
Definition G. 5 (Sobolev spaces). Let $-\infty \leq a<b \leq \infty$ and $1 \leq p \leq \infty$. We define

$$
W^{1, p}(a, b):=\left\{u \in L^{p}(a, b): \exists g \in L^{p}(a, b) \forall \varphi \in \mathscr{D}(a, b): \int_{a}^{b} u \varphi^{\prime}=-\int_{a}^{b} g \varphi\right\}
$$

The space $W^{1, p}(a, b)$ is called (first) Sobolev space. If $p=2$, then we also write $H^{1}(a, b):=W^{1,2}(a, b)$.

By Lemma G.4, the function $g \in L^{p}(a, b)$ is uniquely determined if it exists. In the following, we will write $u^{\prime}:=g$, in accordance with (G.2). We equip $W^{1, p}(a, b)$ with the norm

$$
\|u\|_{W^{1, p}}:=\|u\|_{p}+\left\|u^{\prime}\right\|_{p},
$$

and if $p=2$, then we define the inner product

$$
\langle u, v\rangle_{H^{1}}:=\int_{a}^{b} u v+\int_{a}^{b} u^{\prime} v^{\prime}
$$

which actually yields the norm $\|u\|_{H^{1}}=\left(\|u\|_{2}^{2}+\left\|u^{\prime}\right\|_{2}^{2}\right)^{\frac{1}{2}}$ (which is equivalent to $\left.\|\cdot\|_{W^{1,2}}\right)$.

Lemma G.6. The Sobolev spaces $W^{1, p}(a, b)$ are Banach spaces, which are separable if $p \neq \infty$. The space $H^{1}(a, b)$ is a separable Hilbert space.

Proof. The fact that the $W^{1, p}$ are Banach spaces, or that $H^{1}$ is a Hilbert space, is an exercise. Recall that $L^{p}(a, b)$ is separable (Remark D.46). Hence, the product space $L^{p}(a, b) \times L^{p}(a, b)$ is separable, and also every subspace of this product space is separable. Now consider the linear mapping

$$
T: W^{1, p}(a, b) \rightarrow L^{p}(a, b) \times L^{p}(a, b), \quad u \mapsto\left(u, u^{\prime}\right)
$$

which is bounded and even isometric. Hence, $W^{1, p}$ is isometrically isomomorphic to a subspace of $L^{p} \times L^{p}$ which is separable. Hence $W^{1, p}$ is separable.

Lemma G.7. Let $u \in W^{1, p}(a, b)$ be such that $u^{\prime}=0$. Then $u$ is constant.
Proof. Choose $\psi \in \mathscr{D}(a, b)$ such that $\int_{a}^{b} \psi=1$. Then, for every $\varphi \in \mathscr{D}(a, b)$, the function $\varphi-\left(\int_{a}^{b} \varphi\right) \psi$ is the derivative of a test function since $\int_{a}^{b}\left(\varphi-\left(\int_{a}^{b} \varphi\right) \psi\right)=0$. Hence, by definition,

$$
0=\int_{a}^{b} u\left(\varphi-\left(\int_{a}^{b} \varphi\right) \psi\right)
$$

or, with $c=\int_{a}^{b} u \psi=\mathrm{const}$,

$$
\int_{a}^{b}(u-c) \varphi=0 \quad \forall \varphi \in \mathscr{D}(a, b)
$$

By Lemma G.4, $u=c$ almost everywhere.
Lemma G.8. Let $-\infty<a<b<\infty$ and let $t_{0} \in[a, b]$. Let $g \in L^{p}(a, b)$ and define

$$
u(t):=\int_{t_{0}}^{t} g(s) d s, \quad t \in[a, b] .
$$

Then $u \in W^{1, p}(a, b)$ and $u^{\prime}=g$.

Proof. Let $\varphi \in \mathscr{D}(a, b)$. Then, by Fubini's theorem,

$$
\begin{aligned}
\int_{a}^{b} u \varphi^{\prime} & =\int_{a}^{b} \int_{t_{0}}^{t} g(s) d s \varphi^{\prime}(t) d t \\
& =\int_{a}^{t_{0}} \int_{t_{0}}^{t} g(s) d s \varphi^{\prime}(t) d t+\int_{t_{0}}^{b} \int_{t_{0}}^{t} g(s) d s \varphi^{\prime}(t) d t \\
& =-\int_{a}^{t_{0}} \int_{a}^{s} \varphi^{\prime}(t) d t g(s) d s+\int_{t_{0}}^{b} \int_{s}^{b} \varphi^{\prime}(t) d t g(s) d s \\
& =-\int_{a}^{t_{0}} \varphi(s) g(s) d s-\int_{t_{0}}^{b} \varphi(s) g(s) d s \\
& =-\int_{a}^{b} g \varphi
\end{aligned}
$$

Theorem G.9. Let $u \in W^{1, p}(a, b)$ (bounded or unbounded interval). Then there exists $\tilde{u} \in C((a, b))$ which is continuous up to the boundary of $(a, b)$, which coincides with $u$ almost everywhere and such that for every $s, t \in(a, b)$

$$
\tilde{u}(t)-\tilde{u}(s)=\int_{s}^{t} u^{\prime}(r) d r
$$

Proof. Fix $t_{0} \in(a, b)$ and define $v(t):=\int_{t_{0}}^{t} u^{\prime}(s) d s(t \in \overline{(a, b)})$. Clearly, the function $v$ is continuous. By Lemma G.8, $v \in W^{1, p}(c, d)$ for every bounded interval $(c, d) \subset$ $(a, b)$, and $v^{\prime}=u^{\prime}$. By Lemma G.7, $u-v=C$ for some constant $C$ which clearly does not depend on the choice of the interval $(c, d)$. This proves that $u$ coincides almost everywhere with the continuous function $\tilde{u}=v+C$. By Lemma G.8,

$$
\tilde{u}(t)-\tilde{u}(s)=v(t)-v(s)=\int_{s}^{t} u^{\prime}(r) d r
$$

Remark G.10. By Theorem G.9, we will identify every function $u \in W^{1, p}(a, b)$ with its continuous representant, and we say that every function in $W^{1, p}(a, b)$ is continuous.

Lemma G. 11 (Extension lemma). Let $u \in W^{1, p}(a, b)$. Then there exists $\tilde{u} \in$ $W^{1, p}(\mathbb{R})$ such that $\tilde{u}=u$ on $(a, b)$.

Proof. Assume first that $a$ and $b$ are finite and define

$$
g(t):= \begin{cases}u^{\prime}(t) & \text { if } t \in[a, b] \\ u(a) & \text { if } t \in[a-1, a) \\ -u(b) & \text { if } t \in(b, b+1] \\ 0 & \text { else }\end{cases}
$$

Then $g \in L^{p}(\mathbb{R})$. Let $\tilde{u}(t):=\int_{-\infty}^{t} g(s) d s$, so that $\tilde{u}=u$ on $(a, b)$. By Lemma G.8, $\tilde{u} \in W^{1, p}(c, d)$ for every bounded interval $(c, d) \in \mathbb{R}$. However, $\tilde{u}=0$ outside $(a-$ $1, b+1)$ which implies that $\tilde{u} \in W^{1, p}(\mathbb{R})$.

The case of $a=-\infty$ or $b=\infty$ is treated similarly.
Lemma G.12. For every $1 \leq p<\infty$, the space $\mathscr{D}(\mathbb{R})$ is dense in $W^{1, p}(\mathbb{R})$.
Proof. Let $u \in W^{1, p}(\mathbb{R})$.
Regularization: Choose a positive test function $\varphi \in \mathscr{D}(\mathbb{R})$ such that $\int_{\mathbb{R}} \varphi=1$ and put $\varphi_{n}(x)=n \varphi(n x)$. Then $u_{n}:=u * \varphi_{n} \in C^{\infty} \cap L^{p}(\mathbb{R}), u_{n}^{\prime}=u^{\prime} * \varphi_{n} \in L^{p}(\mathbb{R})$ and

$$
\begin{aligned}
& \lim _{n \rightarrow \infty}\left\|u-u_{n}\right\|_{p}=0 \text { and } \\
& \lim _{n \rightarrow \infty}\left\|u^{\prime}-u_{n}^{\prime}\right\|_{p}=0
\end{aligned}
$$

so that $\lim _{n \rightarrow \infty}\left\|u-u_{n}\right\|_{W^{1, p}}=0$. This proves that $W^{1, p}(\mathbb{R}) \cap C^{\infty}(\mathbb{R})$ is dense in $W^{1, p}(\mathbb{R})$.

Truncation: Choose a sequence $\left(\psi_{n}\right) \subset \mathscr{D}(\mathbb{R})$ such that $0 \leq \psi_{n} \leq 1, \psi_{n}=1$ on $[-n, n]$ and $\left\|\psi_{n}^{\prime}\right\|_{\infty} \leq C$ for all $n \in \mathbb{N}$. Let $\varepsilon>0$. Choose $v \in C^{\infty} \cap W^{1, p}(\mathbb{R})$ such that $\|u-v\|_{W^{1, p}} \leq \varepsilon$ (regularization step). For every $n \in \mathbb{N}$, one has $v \psi_{n} \in \mathscr{D}(\mathbb{R})$ and it is easy to check that for all $n$ large enough, $\left\|v-v \psi_{n}\right\|_{W^{1, p}} \leq \varepsilon$. The claim is proved.
Corollary G.13. For every $u \in W^{1, p}(a, b)$ (bounded or unbounded interval, $1 \leq p<$ $\infty)$ and every $\varepsilon>0$, there exists $v \in \mathscr{D}(\mathbb{R})$ such that $\left\|u-\left.v\right|_{(a, b)}\right\|_{W^{1, p}} \leq \varepsilon$.
Proof. Given $u \in W^{1, p}(a, b)$, we first choose an extension $\tilde{u} \in W^{1, p}(\mathbb{R})$ (extension lemma G.11) and then a test function $v \in \mathscr{D}(\mathbb{R})$ such that $\|\tilde{u}-v\|_{W^{1, p}(\mathbb{R})} \leq \varepsilon$ (Lemma G.12). Then $\|\tilde{u}-v\|_{W^{1, p}(a, b)}=\|u-v\|_{W^{1, p}(a, b)} \leq \varepsilon$.

Corollary G. 14 (Sobolev embedding theorem). Every function $u \in W^{1, p}(a, b)$ is continuous and bounded and there exists a constant $C \geq 0$ such that

$$
\|u\|_{\infty} \leq C\|u\|_{W^{1, p}} \quad \text { for every } u \in W^{1, p}(a, b)
$$

Proof. If $p=\infty$, there is nothing to prove. We first prove the claim for the case $(a, b)=\mathbb{R}$.

So let $1 \leq p<\infty$ and let $v \in \mathscr{D}(\mathbb{R})$. Then $G(v):=|v|^{p-1} v \in C_{c}^{1}(\mathbb{R})$ and $G(v)^{\prime}=$ $p|v|^{p-1} v^{\prime}$. By Hölder's inequality,

$$
|G(v)(x)|=\left.p\left|\int_{-\infty}^{x}\right| v\right|^{p-1} v^{\prime} \mid \leq p\|v\|_{p}^{p-1}\left\|v^{\prime}\right\|_{p}
$$

so that by Young's inequality ( $\left.a b \leq \frac{1}{p} a^{p}+\frac{1}{p^{p^{\prime}}} b^{p^{\prime}}\right)$

$$
\|v\|_{\infty}=\|G(v)\|_{\infty}^{1 / p} \leq C\|v\|_{W^{1, p}} .
$$

Since $\mathscr{D}(\mathbb{R})$ is dense in $W^{1, p}(\mathbb{R})$ by Lemma G.12, the claim for $(a, b)=\mathbb{R}$ follows by an approximation argument.

The case $(a, b) \neq \mathbb{R}$ is an exercise.

Theorem G. 15 (Product rule, partial integration). Let $u$, $v \in W^{1, p}(a, b)(1 \leq p \leq$ $\infty)$. Then:
(i) (Product rule). The product uv belongs to $W^{1, p}(a, b)$ and

$$
(u v)^{\prime}=u^{\prime} v+u v^{\prime}
$$

(ii) (Partial integration). If $-\infty<a<b<\infty$, then

$$
\int_{a}^{b} u^{\prime} v=u(b) v(b)-u(a) v(a)-\int_{a}^{b} u v^{\prime} .
$$

Proof. Since every function in $W^{1, p}(a, b)$ is bounded, we find that $u v, u^{\prime} v+u v^{\prime} \in$ $L^{p}(a, b)$. Choose sequences $\left(u_{n}\right),\left(v_{n}\right) \subset \mathscr{D}(\mathbb{R})$ such that $\left.\lim _{n \rightarrow \infty} u_{n}\right|_{(a, b)}=u$ and $\left.\lim _{n \rightarrow \infty} v_{n}\right|_{(a, b)}=v$ in $W^{1, p}(a, b)$ (Corollary G.13). By Corollary G.14, this implies also $\lim _{n \rightarrow \infty}\left\|\left.u_{n}\right|_{(a, b)}-u\right\|_{\infty}=\lim _{n \rightarrow \infty}\left\|\left.v_{n}\right|_{(a, b)}-v\right\|_{\infty}=0$. The classical product rule implies

$$
\left(u_{n} v_{n}\right)^{\prime}=u_{n}^{\prime} v_{n}+u_{n} v_{n}^{\prime} \text { for every } n \in \mathbb{N}
$$

and the classical rule of partial integration implies

$$
\int_{a}^{b} u_{n}^{\prime} v_{n}=u_{n}(b) v_{n}(b)-u_{n}(a) v_{n}(a)-\int_{a}^{b} u_{n} v_{n}^{\prime} \text { for every } n \in \mathbb{N} .
$$

The claim follows upon letting $n$ tend to $\infty$.
Definition G.16. For every $1 \leq p \leq \infty$ and every $k \geq 2$ we define inductively the Sobolev spaces

$$
W^{k, p}(a, b):=\left\{u \in W^{1, p}(a, b): u^{\prime} \in W^{k-1, p}(a, b)\right\}
$$

which are Banach spaces for the norms

$$
\|u\|_{W^{k, p}}:=\sum_{j=0}^{k}\left\|u^{(j)}\right\|_{p}
$$

We denote $H^{k}(a, b):=W^{k, 2}(a, b)$ which is a Hilbert space for the scalar product

$$
\langle u, v\rangle_{H^{k}}:=\sum_{j=0}^{k} u^{(j)} v^{(j)}{ }_{L^{2}}
$$

Finally, we define

$$
W_{0}^{k, p}(a, b):=\overline{\mathscr{D}(a, b)} \|^{\|\cdot\|_{W^{k, p}}}
$$

that is, $W_{0}^{k, p}(a, b)$ is the closure of the test functions in $W^{k, p}(a, b)$, and we put $H_{0}^{k}(a, b):=W_{0}^{k, 2}(a, b)$.

Theorem G.17. Let $-\infty<a<b<\infty$. A function $u \in W_{0}^{1, p}(a, b)$ if and only if $u \in$ $W^{1, p}(a, b)$ and $u(a)=u(b)=0$.
Theorem G. 18 (Poincaré inequality). Let $-\infty<a<b<\infty$ and $1 \leq p<\infty$. Then there exists a constant $\lambda>0$ such that

$$
\lambda \int_{a}^{b}|u|^{p} \leq \int_{a}^{b}\left|u^{\prime}\right|^{p} \quad \text { for every } u \in W_{0}^{1, p}(a, b)
$$

Proof. Let $u \in W^{1, p}(a, b)$. Then

$$
\begin{aligned}
\int_{a}^{b}|u(x)|^{p} d x & =\int_{a}^{b}\left|\int_{a}^{x} u^{\prime}(y) d y\right|^{p} d x \\
& \leq \int_{a}^{b}\left(\int_{a}^{b}\left|u^{\prime}(y)\right| d y\right)^{p} d x \\
& \leq \int_{a}^{b}(b-a)^{p-1} \int_{a}^{b}\left|u^{\prime}(y)\right|^{p} d y d x \\
& =(b-a)^{p} \int_{a}^{b}\left|u^{\prime}(y)\right|^{p} d y
\end{aligned}
$$

Between the first and the second line, we have used the assumption that $u(a)=0$, while in the following inequality we applied Hölder's inequality.
Theorem G.19. Let $-\infty<a<b<\infty$. For every $f \in L^{2}(a, b)$ there exists a unique function $u \in H_{0}^{1}(a, b) \cap H^{2}(a, b)$ such that

$$
\left\{\begin{array}{l}
u-u^{\prime \prime}=f \quad \text { and }  \tag{G.3}\\
u(a)=u(b)=0
\end{array}\right.
$$

Proof. We first note that if $u \in H_{0}^{1}(a, b) \cap H^{2}(a, b)$ is a solution, then, by partial integration (Theorem G.15), for every $v \in H_{0}^{1}(a, b)$

$$
\begin{equation*}
\int_{a}^{b}\left(u v+u^{\prime} v^{\prime}\right)=(u, v)_{H_{0}^{1}}=\int_{a}^{b} f v . \tag{G.4}
\end{equation*}
$$

By the Cauchy-Schwarz inequality, the linear functional $\varphi \in H_{0}^{1}(a, b)^{\prime}$ defined by $\varphi(v)=\int_{a}^{b} f v$ is bounded:

$$
|\varphi(v)| \leq\|f\|_{2}\|v\|_{2} \leq\|f\|_{2}\|v\|_{H_{0}^{1}} .
$$

By the theorem of Riesz-Fréchet, there exists a unique $u \in H_{0}^{1}(a, b)$ such that (G.4)holds true for all $v \in H_{0}^{1}(a, b)$. This proves uniqueness of a solution of (G.3), and if we prove that in addition $u \in H^{2}(a, b)$, then we prove existence, too. However, (G.4)holds in particular for all $v \in \mathscr{D}(a, b)$, i.e.

$$
\int_{a}^{b} u^{\prime} v^{\prime}=-\int_{a}^{b}(u-f) v \quad \forall v \in \mathscr{D}(a, b)
$$

and $u-f \in L^{2}(a, b)$ by assumption. Hence, by definition, $u^{\prime} \in H^{1}(a, b)$, i.e. $u \in$ $H^{2}(a, b)$ and $u^{\prime \prime}=u-f$. Using also Theorem G.17, the claim is proved.

## G. 3 Sobolev spaces in several dimensions

In order to motivate Sobolev spaces in several space dimensions, we have to recall the partial integration rule in this case (see the Divergence Theorem F.6).

Theorem G. 20 (Gauß). Let $\Omega \subseteq \mathbb{R}^{d}$ be open and bounded such that $\partial \Omega$ is of class $C^{1}$. Then there exists a unique Borel measure $\sigma$ on $\partial \Omega$ such that for every $u, v \in$ $C^{1}(\bar{\Omega})$ and every $1 \leq i \leq d$

$$
\int_{\Omega} u \frac{\partial v}{\partial x_{i}}=\int_{\partial \Omega} u v n_{i} d \sigma-\int_{\Omega} \frac{\partial u}{\partial x_{i}} v
$$

where $n(x)=\left(n_{i}(x)\right)_{1 \leq i \leq d}$ denotes the outer normal vector at a point $x \in \partial \Omega$.
In particular, if $u \in C^{1}(\bar{\Omega})$ and $\varphi \in \mathscr{D}(\Omega)$, then

$$
\int_{\Omega} u \frac{\partial \varphi}{\partial x_{i}}=-\int_{\Omega} \frac{\partial u}{\partial x_{i}} \varphi
$$

Definition G. 21 (Sobolev spaces). Let $\Omega \subseteq \mathbb{R}^{d}$ be any open set and $1 \leq p \leq \infty$. We define

$$
\begin{aligned}
W^{1, p}(\Omega):=\left\{u \in L^{p}(\Omega): \forall 1 \leq i \leq d \exists g_{i}\right. & \in L^{p}(\Omega) \\
& \left.\forall \varphi \in \mathscr{D}(\Omega): \int_{\Omega} u \frac{\partial \varphi}{\partial x_{i}}=-\int_{\Omega} g_{i} \varphi\right\}
\end{aligned}
$$

The space $W^{1, p}(\Omega)$ is called (first) Sobolev space. If $p=2$, then we also write $H^{1}(\Omega):=W^{1,2}(\Omega)$.

Let $u \in W^{1, p}(\Omega)$. By Lemma G.4, the functions $g_{i}$ are uniquely determined. We write $\frac{\partial u}{\partial x_{i}}:=g_{i}$ and call $\frac{\partial u}{\partial x_{i}}$ the partial derivative of $u$ with respect to $x_{i}$. As in the one-dimensional case, the following holds true.

Lemma G.22. The Sobolev spaces $W^{1, p}(\Omega)$ are Banach spaces for the norms

$$
\|u\|_{W^{1, p}}:=\|u\|_{p}+\sum_{i=1}^{d}\left\|\frac{\partial u}{\partial x_{i}}\right\|_{p} \quad(1 \leq p \leq \infty)
$$

and $H^{1}(\Omega)$ is a Hilbert space for the inner product

$$
\langle u, v\rangle_{H^{1}}:=\langle u, v\rangle_{L^{2}}+\sum_{i=1}^{d}\left\langle\frac{\partial u}{\partial x_{i}}, \frac{\partial v}{\partial x_{i}}\right\rangle_{L^{2}} .
$$

## Proof. Exercise.

Not all properties of Sobolev spaces on intervals carry over to Sobolev spaces on open sets $\Omega \subset \mathbb{R}^{d}$. For example, it is not true that every function $u \in W^{1, p}(\Omega)$ is continuous (without any further restrictions on $p$ and $\Omega$ )!
Definition G.23. For every open $\Omega \subset \mathbb{R}^{d}, 1 \leq p \leq \infty$ and every $k \geq 2$ we define inductively the Sobolev spaces

$$
W^{k, p}(\Omega):=\left\{u \in W^{1, p}(\Omega): \forall 1 \leq i \leq d: \frac{\partial u}{\partial x_{i}} \in W^{k-1, p}(\Omega)\right\}
$$

which are Banach spaces for the norms

$$
\|u\|_{W^{k, p}}:=\|u\|_{p}+\sum_{i=0}^{k}\left\|\frac{\partial u}{\partial x_{i}}\right\|_{W^{k-1, p}}
$$

We denote $H^{k}(\Omega):=W^{k, 2}(\Omega)$ which is a Hilbert space for the inner product

$$
\langle u, v\rangle_{H^{k}}:=\langle u, v\rangle_{L^{2}}+\sum_{i=0}^{k}\left\langle\frac{\partial u}{\partial x_{i}}, \frac{\partial v}{\partial x_{i}}\right\rangle_{H^{k-1}}
$$

Finally, we define

$$
W_{0}^{k, p}(\Omega):=\overline{\mathscr{D}(\Omega)}{ }^{\|\cdot\|_{W^{k}, p}},
$$

that is, $W_{0}^{k, p}(\Omega)$ is the closure of the test functions in $W^{k, p}(\Omega)$, and we put $H_{0}^{k}(\Omega):=W_{0}^{k, 2}(\Omega)$.
Theorem G. 24 (Poincaré inequality). Let $\Omega \subset \mathbb{R}^{d}$ be a bounded domain, and let $1 \leq p<\infty$. Then there exists a constant $C \geq 0$ such that

$$
\int_{\Omega}|u|^{p} \leq C^{p} \int_{\Omega}|\nabla u|^{p} \quad \text { for every } u \in W_{0}^{1, p}(\Omega)
$$

We note that the Poincaré inequality implies that

$$
\|u\|:=\left(\int_{\Omega}|\nabla u|^{p}\right)^{\frac{1}{p}}
$$

defines an equivalent norm on $W_{0}^{1, p}(\Omega)$ if $\Omega \subset \mathbb{R}^{d}$ is bounded. Clearly,

$$
\|u\| \leq\|u\|_{W_{0}^{1, p}} \quad \text { for every } u \in W_{0}^{1, p}
$$

by the definition of the norm in $W^{1, p}$. On the other hand,

$$
\begin{aligned}
\|u\|_{W_{0}^{1, p}} & \leq C\left(\|u\|_{L^{p}}+\|\nabla u\|_{L^{p}}\right) \\
& \leq C\|\nabla u\|_{L^{p}}=C\|u\|,
\end{aligned}
$$

by the Poincaré inequality.
We also state the following two theorems without proof.
Theorem G. 25 (Sobolev embedding theorem). Let $\Omega \subset \mathbb{R}^{d}$ be an open set with $C^{1}$ boundary. Let $1 \leq p \leq \infty$ and define

$$
p^{*}:= \begin{cases}\frac{d p}{d-p} & \text { if } 1 \leq p<d \\ \infty & \text { if } d<p\end{cases}
$$

and if $p=d$, then $p^{*} \in[1, \infty)$. Then, for every $p \leq q \leq p^{*}$ we have

$$
W^{1, p}(\Omega) \subset L^{q}(\Omega)
$$

with continuous embedding, that is, there exists $C=C(p, q) \geq 0$ such that

$$
\|u\|_{L^{q}} \leq C\|u\|_{W^{1, p}} \quad \text { for every } u \in W^{1, p}(\Omega) .
$$

Theorem G. 26 (Rellich-Kondrachov). Let $\Omega \subset \mathbb{R}^{d}$ be an open and bounded set with $C^{1}$ boundary. Let $1 \leq p \leq \infty$ and define $p^{*}$ as in the Sobolev embedding theorem. Then, for every $p \leq q<\infty$ the embedding

$$
W^{1, p}(\Omega) \subset L^{q}(\Omega)
$$

is compact, that is, every bounded sequence in $W^{1, p}(\Omega)$ has a subsequence which converges in $L^{q}(\Omega)$.

## G. 4 * Elliptic partial differential equations

Let $\Omega \subset \mathbb{R}^{d}$ be an open, bounded set, $f \in L^{2}(\Omega)$, and consider the elliptic partial differential equation

$$
\begin{cases}u-\Delta u=f & \text { in } \Omega  \tag{G.5}\\ u=0 & \text { in } \partial \Omega\end{cases}
$$

where

$$
\Delta u(x):=\sum_{i=1}^{d} \frac{\partial^{2}}{\partial x_{i}^{2}} u(x)
$$

stands for the Laplace operator.
If $u \in H_{0}^{1}(\Omega) \cap H^{2}(\Omega)$ is a solution of (G.5), then, by definition of the Sobolev spaces, for every $v \in \mathscr{D}(a, b)$

$$
\begin{aligned}
\langle u, v\rangle_{H_{0}^{1}} & =\int_{\Omega}\left(u v+\sum_{i=1}^{d} \frac{\partial u}{\partial x_{i}} \frac{\partial v}{\partial x_{i}}\right) \\
& =\int_{\Omega}\left(u v-\sum_{i=1}^{d} \frac{\partial^{2} u}{\partial x_{i}^{2}} v\right) \\
& =\int_{\Omega}(u-\Delta u) v \\
& =\int_{\Omega} f v .
\end{aligned}
$$

By density of the test functions in $H_{0}^{1}(\Omega)$, this equality holds actually for all $v \in$ $H_{0}^{1}(\Omega)$. This may justify the following definition of a weak solution.

Definition G.27. A function $u \in H_{0}^{1}(\Omega)$ is called a weak solution of (G.5) if for every $v \in H_{0}^{1}(\Omega)$

$$
\begin{equation*}
\langle u, v\rangle_{H_{0}^{1}}=\int_{\Omega} u v+\int_{\Omega} \nabla u \nabla v=\int_{\Omega} f v, \tag{G.6}
\end{equation*}
$$

where $\nabla u$ is the usual, euclidean gradient of $u$.
Theorem G.28. Let $\Omega \subset \mathbb{R}^{d}$ be an open, bounded set. Then, for every $f \in L^{2}(\Omega)$ there exists a unique weak solution $u \in H_{0}^{1}(\Omega)$ of the problem (G.5).

Proof. By the Cauchy-Schwarz inequality, the linear functional $\varphi \in H_{0}^{1}(\Omega)^{\prime}$ defined by $\varphi(v)=\int_{\Omega} f v$ is bounded:

$$
|\varphi(v)| \leq\|f\|_{2}\|v\|_{2} \leq\|f\|_{2}\|v\|_{H_{0}^{1}} .
$$

By the theorem of Riesz-Fréchet, there exists a unique $u \in H_{0}^{1}(\Omega)$ such that (G.6) holds true for all $v \in H_{0}^{1}(a, b)$. The claim is proved.


[^0]:    ${ }^{1}$ To tell the truth, by geometrical arguments, Huygens disproved Galilei's conjecture before the discovery of differential calculus and in addition solved the tautochrone problem. Only later, by using differential calculus, Lagrange, Euler and Abel gave analytic proofs of the fact that the cycloid is the solution to this problem.

[^1]:    ${ }^{2}$ The original citation from the 1948 edition of his book on Functional Analysis is: I hail a semigroup whenever I see one and I seem to see them everywhere. This vision has been largely shared by many respectful mathematicians.

[^2]:    ${ }^{1}$ to dissipate (lat.: dissipare): to cause to lose energy (such as heat) irreversibly (http: / / www . thefreedictionary.com/). Other explanations are: to spend or expend intemperately or wastefully. Or: to attenuate to or almost to the point of disappearing. Of course, a constant function is an energy function for every ordinary differential equation, but this trivial energy function does not allow us to call every differential equation dissipative. One should have the additional property from Lemma 1.5 in mind.

[^3]:    ${ }^{2}$ This idea has already been proposed by Augustin Cauchy in a short communication to the Comptes Rendus de l'Académie des Sciences de Paris in 1847, Méthodes générale pour la résolution des systèmes d'équations simultanées, volume 25 (1847), pages 536-538.

[^4]:    ${ }^{1}$ See D. T. Whiteside (ed.), The Mathematical Papers of Isaac Newton, Volume II, 1667-1670, Cambridge University Press, Cambridge, 1968 and D. T. Whiteside (ed.), The Mathematical Papers of Isaac Newton, Volume III, 1670-1673, Cambridge University Press, Cambridge, 1969.
    ${ }^{2}$ The observation that the continuous Newton method is a gradient system is taken from J. W. Neuberger, Prospects for a central theory of partial differential equations, Math. Intelligencer 27 (2005), 47-55

[^5]:    ${ }^{1}$ Oeuvres complètes de Laplace. Tome 10 / publiées sous les auspices de l'Académie des sciences, par MM. les secrétaires perpétuels - Gauthier-Villars (Paris) - 1878-1912. See the memoir Théorie des attractions des sphéroüdes et de la figure des planètes on page 341. It is for us difficult to translate the beginning of the citation, but a translation could be: ... and I dare to be proud to present to the geometers, in this work, a theory of attractions of spheroids and of the shape of the planets which is more general and simpler than those which are already known.
    ${ }^{2}$ Op. cit., pages 341-.

[^6]:    ${ }^{3}$ Jahnke, H. N. (Ed.). : A history of analysis. Vol. 24 of History of Mathematics. American Mathematical Society, Providence, RI, 2003, p. 198.
    ${ }^{4}$ Oeuvres complètes de Laplace. Tome 10 . Op. cit., page 362 . We denote the spherical coordinates slightly differently.
    ${ }^{5}$ Oeuvres complètes de Laplace. Tome 11 / publiées sous les auspices de l'Académie des sciences, par MM. les secrétaires perpétuels - Gauthier-Villars (Paris) - 1878-1912. See the Mémoire sur la théorie de l'anneau de Saturne.

[^7]:    ${ }^{6}$ Oeuvres complètes de Laplace. Tome 11. Op. cit., pages 277-278.
    ${ }^{7}$ The memoir Théorie des attractions ... was submitted to the Académie des Sciences in 1782, and published in 1785.

[^8]:    ${ }^{1}$ autonomous (greek: auto+nomos, nomos = law; http://www.thefreedictionary. $\mathrm{com} /$ ): not controlled by others or by outside forces; independent. Or: Independent of the laws of another state or government; self-governing.

[^9]:    ${ }^{1}$ Albert Einstein, Geometrie und Erfahrung (Geometry and experience), An Address to the Prussian Academy of Sciences in Berlin on January 27th, 1921. German original quotation: Insofern sich die Sätze der Mathematik auf die Wirklichkeit beziehen, sind sie nicht sicher, und insofern sie sicher sind, beziehen sie sich nicht auf die Wirklichkeit.

[^10]:    ${ }^{1}$ The letters $\alpha$ and $\omega$ are the first and the last letter, respectively, of the greek alphabet. While the " $\omega$-limit set" gives information about the behaviour of a continuous function near $+\infty$, the " $\alpha$-limit set" is defined analogously for a continuous function defined on $\mathbb{R}_{-}$and gives information about the behaviour near $-\infty$.

[^11]:    ${ }^{2}$ The citation is taken from H. B. Curry, The method of steepest descent for non-linear minimization problems, Quart. Appl. Math. 2 (1944). We have changed the notation for the energy $\mathscr{E}$ which was denoted by $G$ in the original.

